Explor Anim Med Res,
Vol.10, Issue - 2, 2020, p. 169-178

ISSN 2277- 470X (Print), ISSN 2319-247X (Online)
Website: www.animalmedicalresearch.org

Research Article

EVOLUTIONARY ANALYSIS AND DETECTION OF POSITIVE SELECTION OF HEMAGGLUTININ AND NEURAMINIDASE GENES OF H5N1 AVIAN INFLUENZA FROM CHICKEN, DUCK AND GOOSE ACROSS ASIA

Kuralayanapalya Puttahonnappa Suresh*, Sharanagouda Patil, Uma Bharathi Indrabalan, Rajangam Sridevi, Paramanadham Krishnamoorthy, Shinduja Rajamani, Parimal Roy

Received 15 September 2020, revised 23 November 2020

ABSTRACT: Outbreaks of very high pathogenic avian influenza (H5N1) viruses are being reported in poultry in almost all countries including Asia. It has been reported that the spread is very fast and found that this virus is spreading in avian species since several years. In this study, the evidence of positive selection prominent to mutations was analyzed for the Hemagglutinin (HA) and Neuraminidase (NA) nucleotide sequences of H5N1 avian influenza from chicken, duck and goose across Asia. H5N1 avian influenza viruses are being a severe risk to the public health. Detection of positive selection sites in Hemagglutinin (HA) and Neuraminidase (NA) genes will help to trace the evolutionary path of these viruses from different poultry hosts. The positive/ diversifying selection (*dN/dS* (ω) >1) was found to be showing significant signals in mutation of HA and NA genes and is evolving rapidly. The cumulative *dN/dS* (ω) ratio was found ranging from 0.21 to 0.23 in HA gene and 0.16 to 0.25 in NA gene of Avian Influenza Virus from chicken, duck and goose. Furthermore, statistical Bayesian model methods were applied to interpret the genetic diversity of H5N1 strain, the evolutionary rates were ranging from $2.36x10^{-3}$ to $5.19x10^{-3}$ in HA gene and $2.28x10^{-3}$ to $6.25x10^{-3}$ in NA gene from chicken, duck and goose respectively, which revealed a rapid evolution in these viruses with respect to their genetic ancestor. Substitution rates and selection pressure in these three different hosts indicate that their dynamics of mutation and replication remain similar among the species studied and are important for evolution.

Key words: Avian influenza virus, Bayesian analysis, Chicken, Duck, Goose, Hemagglutinin, Neuraminidase, Positive selection.

INTRODUCTION

Avian influenza viruses are single-stranded negative-sense Ribonucleic acid (RNA) belongs to the family *Orthomyxoviridae*, its genome consists of eight segments. The segmented nature of this viral genome which has polymerase prone to error that enhances adaptation in new hosts and rapid evolution of the virus (Klen *et al.* 2007). Influenza A viruses, including avian influenza viruses (AIVs), are enveloped, pleomorphic, and eight segments ranging from 890 to 2341 nucleotides in length (Webster *et al.* 1992). These viruses cause mild to severe infection in different bird species and are designated as low-morbific to highly morbific avian influenza viruses (Suartha *et al.* 2018).

There are four types of influenza viruses: A, B, C and D, and predominantly ducks, geese and chickens are the natural hosts for most influenza type A virus.

Type A (H5N1) avian influenza viruses are subtyped centered on two glycoproteins *i.e.* Hemagglutinin (HA) and Neuraminidase (NA) seen on their exterior surface. There are 18 HA and 11 NA subtypes identified so far. H5N1 virus has HA5 protein and an NA1 protein, used in our study. They interfere in the host cell attachments and has the transmitting ability to cause influenza virus disease to spread rapidly. HA binds to the sialic acid of monosaccharide present on the exterior surface of the host target cells, whereas the function of neuraminidase is to remove receptors for the influenza virus from newly formed virus particles allowing these to be released and spread the infection (Pushko *et al.* 2017). Sialic acid of monosaccharide involves restraints for the virus, which is destroyed by the Neuraminidase gene. An outbreak of

ICAR-National Institute of Veterinary Epidemiology and Disease Informatics (NIVEDI), Yelahanka, Bengaluru-560064, Karnataka, India.
**Corresponding author. e-mail: sureshkp97@rediffmail.com*

H5N1 spreading regionally has been transmitted among birds in Asian countries (Tian *et al.* 2015).

These viruses naturally occur amongst wild birds are also found to infect domestic birds and other avian species. Infectious type A, (H5N1) avian influenza virus causes huge mortality amongst the avian species. Avian influenza viruses are RNA viruses that replicate very rapidly among the poultry birds.

HA and NA genes undergo mutations rapidly and highly adaptable to evolve themselves to survive in the host. In this regard, study was undertaken to estimate the evolutionary rate, time of the most recent common ancestor (tMRCA) and positive selection sites to delineate the evolutionary pathway of virus. Molecular evolutionary analysis tool such as Bayesian Evolutionary Analysis by Sampling Trees (BEAST) (Hassan *et al.* 2020) which estimates genetic variation obtained in this regard in these datasets to get the evolutionary rates.

In this paper, we evaluated and estimated the ability of standard modelling tools to obtain the nucleotide substitution rate and tMRCA for the data downloaded from the Influenza research database wherein the evolutionary dates are taken as sample dates (Poen *et al.* 2019).

The similarity of substitution rates with non-synonymous to synonymous values is the approach most commonly used to estimate positive selection. Positive selection was interpreted using (*dN*/*dS or ω)* ratio in the context of codons that constitutes the ratio of the non-synonymous rate (*dN*) to synonymous rates (*dS*) of evolution. Interpretations are usually with instance to non-synonymous/synonymous (*ω*) substitution rate ratio per region (Mahardikaa *et al.* 2016).

Wherein the rates *dN* and *dS* are the numbers substitutions of nonsynonymous and synonymous rates per site, respectively. The *dN/dS* (*ω*) ratio estimates the selection pressure at each protein-coding (El-Shesheny *et al.* 2016). The ratio $\omega > 1$ indicates positive selection, $\omega = 1$ indicates neutral selection and $0 > \omega > 1$ indicates negative selection.

The synonymous mutation rate is used as a base rate of neutral evolution while comparing to the effect of the non-synonymous substitution, the selective effect of a synonymous substitution is insignificant. A fixed-effects likelihood (FEL) method measures non-synonymous and synonymous rates of substitution at each codon site. FEL describes the model of rate variation over improved random effects models, which does not experience much false positives as random effects models do for data sets. Single Likelihood Ancestor Counting (SLAC) (Arafa *et al.* 2016) provides inheritance functionality which adapts counting-based method for phylogenetic applications (Lee *et al.* 2016).

The procedure for model test is conducted on likelihood ratio tests repeatedly between nested model, and models of non-nested for comparisons of Akaike's Information Criterion (AIC). A goodness-of-fit criterion Akaike's Information Criterion rewarding the model for higher log likelihood score log(L) and AIC = -2logL+2p. The lowest AIC model explains the best fit for data. The main aim of this study is to obtain the tMRCA and positive selection pressure of HA and NA genes of H5N1 avian influenza viruses from chicken, duck and goose.

MATERIALS AND METHODS

Sequence data

Full-length nucleotide sequences encoding proteins (HA and NA) of type A, Avian Influenza Virus (AIV) from the Influenza Research Database (IRD) were downloaded (http://www.fludb.org) (Zhang *et al.* 2017). Three distinct data sets were generated for each gene centered on their subtype H5N1 from chicken, duck and goose.

Sequence alignment

Multiple sequence alignments were performed using Multiple Alignment using Fast Fourier Transform (MAFFT) sequence alignment tool and Gblocks a biological sequence alignment editor was used for sequence analysis and editing (Castresana *et al.* 2000, Katoh *et al.* 2019).

Homologous recombination detection

RDP4 tool along with the seven recombination detection methods, Rescan/Bootscan (Martin *et al.* 2005), GENECOV (Padidam *et al.* 1999), MaxChi (Smith 1992), RDP (Martin and Rybicki, 2000), 3Seq (Lam *et al.* 2018), Siscan (Gibbs *et al.* 2000) and the Chimaera (Posada and Crandall, 2001) methods detected recombination signals (Darriba *et al.* 2012). RDP4 v.4.100 tool was used for the sequence data (HA and NA) which performs a number of recombination detection and analyzing the techniques.

Evolutionary rate analysis

Based on Akaike Information Criteria (AIC), models were selected using the tool jModelTest2 (Rambaut *et al.* 2016). TempEst v1.5.3 tool was used to evaluate the temporal signals. Using the isolation year as sampling time, BEAUti was used for input analysis (Drummond and Rambaut, 2007). Four molecular clock models (strict clock, Relaxed Clock log normal (RCLN), relaxed clock exponential (RCE) and Random Local Clock (RLC))

Table 1. List of countries that reported avian influenza during 1996-2019.

HA_Goose			NA_Goose		
Accession Number	**Year of collection**	**Country of origin**	**Accession Number**	**Year of collection**	**Country of origin**
MK616077	2015	China	MK616079	2015	China
MF116310	2012	China	MF116320	2012	China
KC261464	2008	China	KC261474	2008	China
DQ997405	2003	China	DQ997406	2003	China
AF144305	1996	China	AF144304	1996	China
AF148678	1996	China	AF364335	1997	China
AF364334	1997	China	HM172181	2004	China
HM172089	2004	China	DQ997523	2001	China
DQ997522	2001	China	HM172168	2005	China
HM172070	2005	China	KR010413	2014	China
KR010411	2014	China	JQ638678	2010	China
JQ638674	2010	China	DQ997277	2003	China
DQ997276	2003	China	GU252832	2008	India
GU252830	2008	India	DQ366316	2005	Vietnam
DQ366314	2005	Vietnam	LC500368	2018	Vietnam
LC500366	2018	Vietnam	LC364291	2017	Vietnam
LC364289	2017	Vietnam	KY437769	2014	China
KY437767	2014	China	CY030907	2006	China
CY030905	2006	China	CY030915	2006	China
CY030913	2006	China	KU042801	2014	China
KU042743	2014	China	KU042802	2014	China
KU042744	2014	China			

were used along with four demographic model (constant population, exponential population, Bayesian skyline plot, extended Bayesian skyline plot) and approximations of nucleotide substitution rates were accomplished using the Markov Chain Monte Carlo (MCMC) method.

Each data with approximate MCMC chains of 30-60 million generations were run and to achieve convergence in each case, chain lengths were run sufficiently and when employed in BEAST tool as burn-in a 10% of samples were discarded. The tMRCA with 95% Highest Posterior Density (HPD) intervals of the substitution rate and divergence time parameters were calculated from the treeheight/tmrca. Mean clock rate and UCLD mean sample respectively. The log files obtained from Beast tool were visualized and the models were paralleled using Bayes factor with approximate values implemented in Tracer v1.7.1 software tool for selection.

A total of six phylogenetic trees, three each for HA (Chicken, duck and goose) and NA (chicken, duck and goose) were processed by tree annotator, which are then visualized with Figtree software (Rambaut *et al.* 2018, Weaver *et al.* 2018).

Positive selection

Datamonkey Adaptive Evolution (DAE) server is a work process that removes the redundant sequences before the analysis (Pond and Frost 2005). The similarity from substitution rates of non-synonymous to synonymous is the approach most commonly used to estimate positive selection. Positive selection is interpreted using the ratio of non-synonymous (*dN*) rate to synonymous rates (*dS*) in the context of codons. The *dN*, *dS* and *dN/dS(ω)* values at each nucleotide site is estimated by the Fixed-Effects Likelihood (FEL), combined with maximum likelihood at every codon-site that fits a codon model (for 20-40 nucleotide sequences). Such codon model using the χ^2 distribution tests the significance of ω value. Single Likelihood Ancestor Counting (SLAC) combined with maximum likelihood (for more than 40 nucleotide sequences) reconstructs

Table 2. Values of substitution rates and tMRCA.

Gene	Host	Substitution Rate (10^{-3}) (subs/site/year)				tMRCA			
				95% HPD				95% HPD	
		Mean	Median	Lower	Upper	Mean	Median	Lower	Upper
HA	Chicken	2.36	2.31	1.82	3.00	69.53	67.63	59.39	83.1
	Duck	5.15	5.15	4.75	5.54	25.22	25.19	23.35	27.16
	Goose	5.19	5.19	4.52	5.89	23.87	23.74	22.62	25.4
NA	Chicken	2.88	3.00	1.82	3.42	36.23	35.14	30	46.61
	Duck	2.28	2.16	1.80	2.98	41.27	40.67	35.71	48.6
	Goose	6.25	6.24	5.39	7.13	22.15	22.1	22	22.47

*tMRCA: time of the most common ancestor in years, HPD: highest posterior density.

ancestral state and estimates *dN*, *dS* and *dN/dS (ω)* (Fourment and Holmes 2015). The *ω* values were estimated based on relative GTR branch lengths and nucleotide substitution biases with Log (L) and AIC-c values recorded for different data sets for Hemagglutinin (HA) and Neuraminidase (NA) gene from the avian hosts chicken, duck and goose.

RESULTS AND DISCUSSION

Dataset

Full-length nucleotide sequences encoding proteins (HA and NA) of type A Avian Influenza Virus (AIV) from 1996-2019 from the Influenza Research Database (IRD) were downloaded (http://www.fludb.org). Three distinct data sets were generated for each gene centered with subtype H5N1. The H5N1 nucleotide sequence from datasets of Hemagglutinin (HA) gene originating from Chicken (n=364), Duck (n=296), Goose (n=22) and Neuraminidase (NA) gene sequences from Chicken (n=361), Duck (n=296), Goose (n=21) were extracted. Table 1 contains the list of countries that reported avian influenza of goose (HA and NA) during 1996-2019.

Sequences alignment

Three distinct nucleotide sequence datasets were generated for HA gene and NA gene from three avian hosts (chicken, goose and duck) centered on their subtype (H5N1) through the web site IRD in FASTA format. The generated nucleotide sequence datasets for Hemagglutinin (HA) gene from avian hosts viz., Chicken (n=364; 1707 bp in length), Duck (n=296; 1704 bp in length), Goose (n=22; 1704 bp in length) and Neuraminidase (NA) gene of Chicken (n=361; 1329 bp in length), Duck (n=296; 1350 bp in length), Goose (n=21; 1350 bp in length). Multiple aligned sequences were obtained from the tool MAFFT and edited by Gblocks and were then saved in FASTA subjected to homologous recombination detection.

Homologous recombination detection

RDP4 v.4.100 tool was employed to undertake a complete exploratory scan on the dataset containing sequences to detect the recombination events. Recombination regions were detected in HA gene specific to hosts chicken and duck and are removed. However, no recombination regions were detected in HA gene specific for goose. The recombination regions for NA gene specific for were detected and are removed. All the edited sequence datasets without recombinant regions were then saved in FASTA format for evolutionary and positive selection analysis.

Evolutionary rate analysis

To know whether there are substantial transformations in evolutionary rates amongst AIV from chicken, duck and goose, the study was focused on HA and NA proteins. Complete gene sequences of HA & NA were used to conclude the substitution rate and time of most common ancestors (tMRCAs) using the Bayesian-based coalescent methodology. jModelTest2 tool was employed to arrive a best fit model as GTR. The data, priors, clock rate and MCMC chain lengths were edited and saved using BEAUti tool in XML format of BEAST tool. The logarithmic (log) and trees files for each HA and NA gene of H5N1 virus from three hosts were obtained from the XML file after executing in BEAST tool. The log files were used to estimate the evolutionary rates and tMRCA. The individual datasets for demographic model were

Table 3. List of positively/diversifying selected sites and cumulative *dN/dS* ratio.

Gene	Host	Method used	Codon Sites	Length of gene (bp)	Positive/diversifying selected sites	dN/dS at positive selected sites	Cumulative *dN/dS*
HA	Chicken	SLAC	568	1707	87, 102, 137, 145, 154, 156, 157, 171	2.05,2.00,2.03,1.04,1.17,2.60,2.00,2.74	0.2357
	Duck	SLAC	567	1704	156, 170, 171, 172, 204, 205	2.16,2.68,2.72,2.00,2.31,2.16	0.2187
	Goose	FEL	567	1704	87, 131, 157, 156, 171	2.79,2.80,3.07,3.87,2.77	0.2185
NA	Chicken	SLAC	442	1329	16, 17, 38, 60, 235	2.16,2.25,2.34,2.12,2.66	0.2302
	Duck	SLAC	449	1350	44, 46, 48, 56, 64	2.64,2.00,1.99,2.00,2.09	0.25
	Goose	FEL	449	1350	44, 46, 369	3.32,2.85,2.78	0.1694

*Positively selected codon with with p-value threshold of 0.1 are reported.

selected with 95% of the Highest Posterior Density (HPD) estimates. Such selected individual HA and NA datasets were used to estimate the tMRCA and substitution rate per site per year by employing RCLM and URCM clock models. Approximations of nucleotide substitution rate were accomplished using the Markov chain Monte Carlo (MCMC) method available in BEAST v1.8.4.

The evolutionary rates for the HA gene from chicken, duck and goose are $2.36x10^{-3}$,$5.15\ x10^{-3}$ and $5.19\ x10^{-3}$ respectively, whereas the evolutionary rates for NA gene from chicken, duck and goose are $2.88x10^{-3}$, $2.28x10^{-3}$ and $6.25x10^{-3}$ respectively. Fig.1 shows the evolutionary rates plotted for both the HA and NA genes, in both HA and NA genes goose shows lineage among other birds taken in this study (chicken, duck and goose).

The tMRCA recorded in ages for the HA gene of chicken as 69.53 years with 95% HPD range (59.39 to 83.10), duck as 25.22 years with 95% HPD range (23.35 to 27.16), goose as 23.87 years with 95% HPD range (22.62 to 25.40) and for NA gene of chicken as 36.23 years with 95% HPD range (30.00 to 46.61), duck as 41.27 years with 95% HPD range (35.71 to 48.60), goose as 22.15 years with 95% HPD range (22.00 to 22.47) represented in Table 2. The HA gene from the host chicken, duck and goose has revealed of 1979, 1995 and 1984 years as its tMRCA respectively, whereas that of NA gene from chicken, duck and goose revealed 1952, 1988 and 1996 years as its tMRCA respectively.

Positive Selection Analysis

The Analysis of HA and NA genes obtained in this study revealed that some of the codon sites had undergone positive selection (ω >1). The number of positively selected sites for glycoprotein HA gene of H5N1 from chicken were in 8 positions (87,102,137,145,154,156,157 and 171) with *dN/dS* (2.05,2.00,2.03,1.04,1.17,2.60,2.00,2.74), from duck were in 6 positions (156,170,171,172,204 and 205) with *dN/dS* (2.16,2.68,2.72,2.00,2.31,2.16) from duck, and from goose were in 5 positions (87,131,157,156,171) with *dN/dS* (2.79,2.80,3.07,3.87,2.77). The cumulative ω values for HA gene from chicken, duck and goose are 0.2357, 0.2187 and 0.2185 respectively. The number of positively selected sites for NA gene of chicken were at 5 positions (16,17,38,60,235) with *dN/dS* (2.16,2.25,2.34,2.12,2.66), duck at 5 positions (44,46,48,56,64) with *dN/dS* (2.64,2.00,1.99,2.00,2.09) and goose at 3 positions (44,46,369) with *dN/dS* (3.32,2.85,2.78). The cumulative ù values for NA gene of chicken duck, and goose were 0.2302, 0.2500 and 0.1694 respectively represented in Table 3. The ù values at each codon site for each host of HA and NA gene were graphically visualized (Fig.2 - Fig.7).

The substitution rates varied between $2.36x10^{-3}$ and $5.19x10^{-3}$ subs/site/year for HA gene among the three datasets from chicken, duck and goose, whereas the substitution rates varied between $2.88x10^{-3}$ and $6.25x10^{-3}$ subs/site/year for NA gene among all the three host species. It means that the divergence of HA gene occurred earlier than that of NA gene in all three avian hosts (Nonthabenjawan *et al.* 2016). The tMRCA in ages for HA from the three hosts were 69.53 years (1997-2019), 25.22 years (2001-2018) and 23.87 years (1996-2018) which shows that the HA gene of AIV from chicken

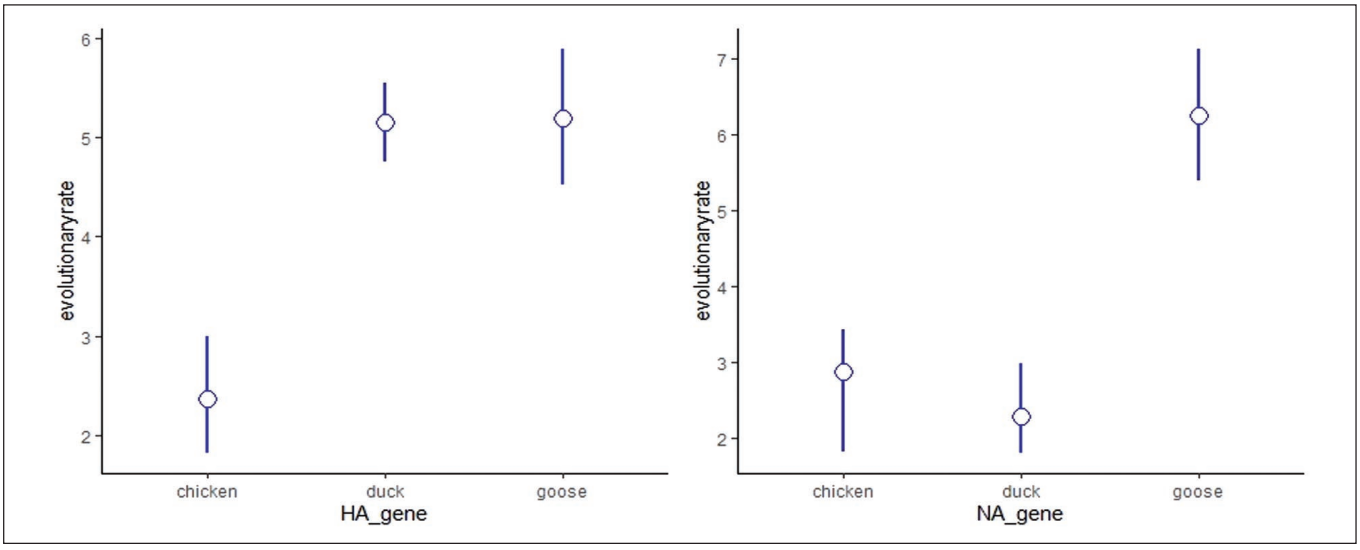

Fig. 1. Evolutionary rates 10^{-3} (subs/site/year) of avian influenza virus of HA and NA gene of chicken, duck and goose 95% higher and lower HPD was evaluated using BEAST.

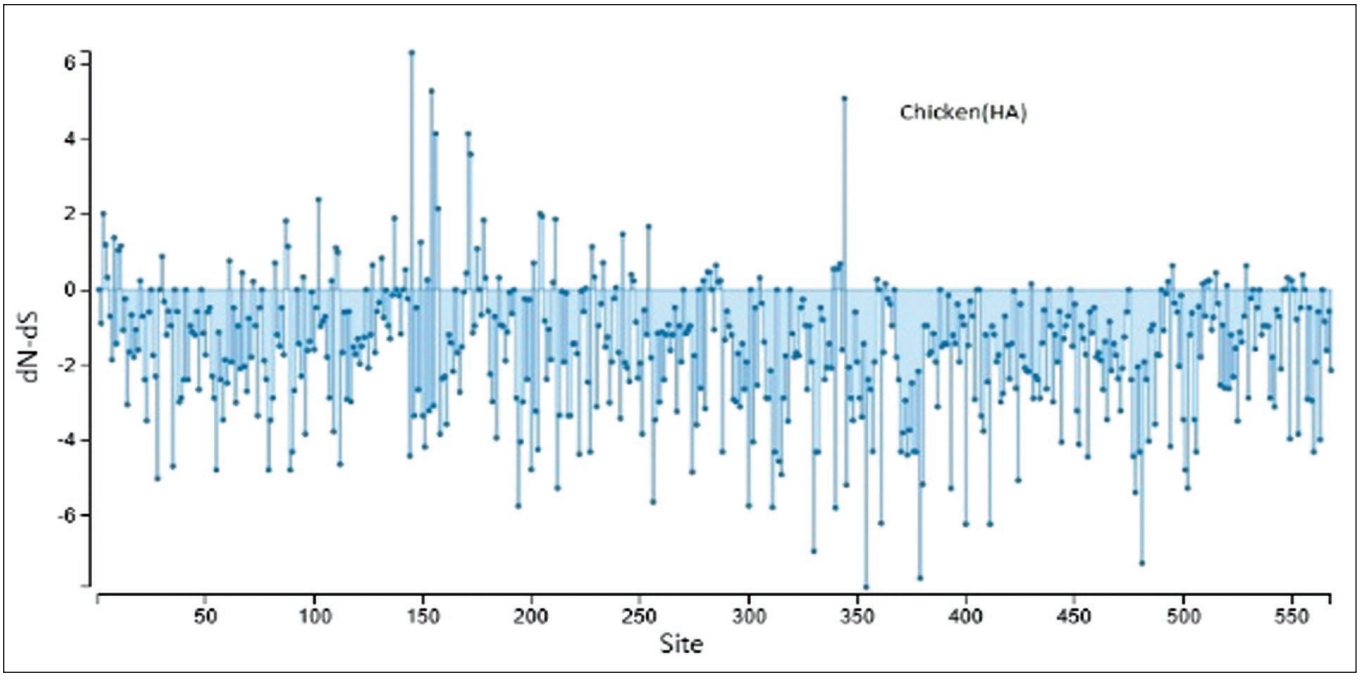

Fig. 2. The positive substitutions per site based on a set of codon aligned nucleotide sequences of chicken H5N1 virus. Cumulated dN-dS are displayed for the codon sites from 1 to 568.

evolved late followed by that from duck and goose. The NA gene tMRCA in ages were 36.23 years (1997-2019), 41.27 years (2001-2018) and 22.15 years (1996-2018) from the three avian hosts respectively, which revealed the evolution of NA gene from goose was earlier, followed by chicken and duck. The nucleotide substitution rate / site/year for the HA and NA genes of H5N1 virus from three different avian host species were estimated by Bayesian analysis with BEAST tool which showed that these viruses were rapidly evolving and transmitting in the host that aggravated faster evolution (Li *et al.* 2011).

In addition to evolutionary rate, nucleotide substitution rates also played an important role in virus evolution. It was found that the selection pressure acting on the HA and NA genes were an important factor for mutation. Hemagglutinin (HA) and Neuraminidase (NA) genes are major glycoproteins placed on external surface of virus interacting with host immune systems frequently. So in the process of positive selection, increased selective pressure may result in more accumulation of phenotypic variants, some of them may evolve as more pathogenic variants. Positively selected variants will be more

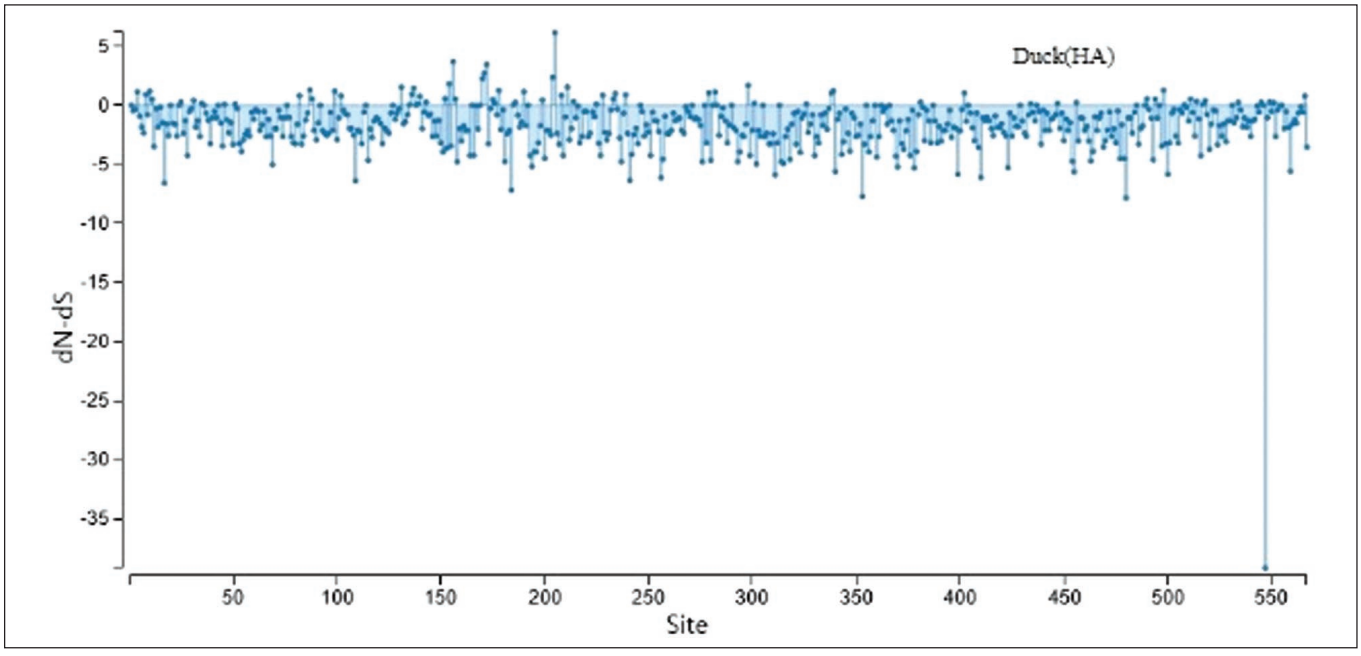

Fig. 3. The positive substitutions per site based on a set of codon aligned nucleotide sequences of duck H5N1virus. Cumulated dN-dS are displayed for the codon sites from 1 to 567.

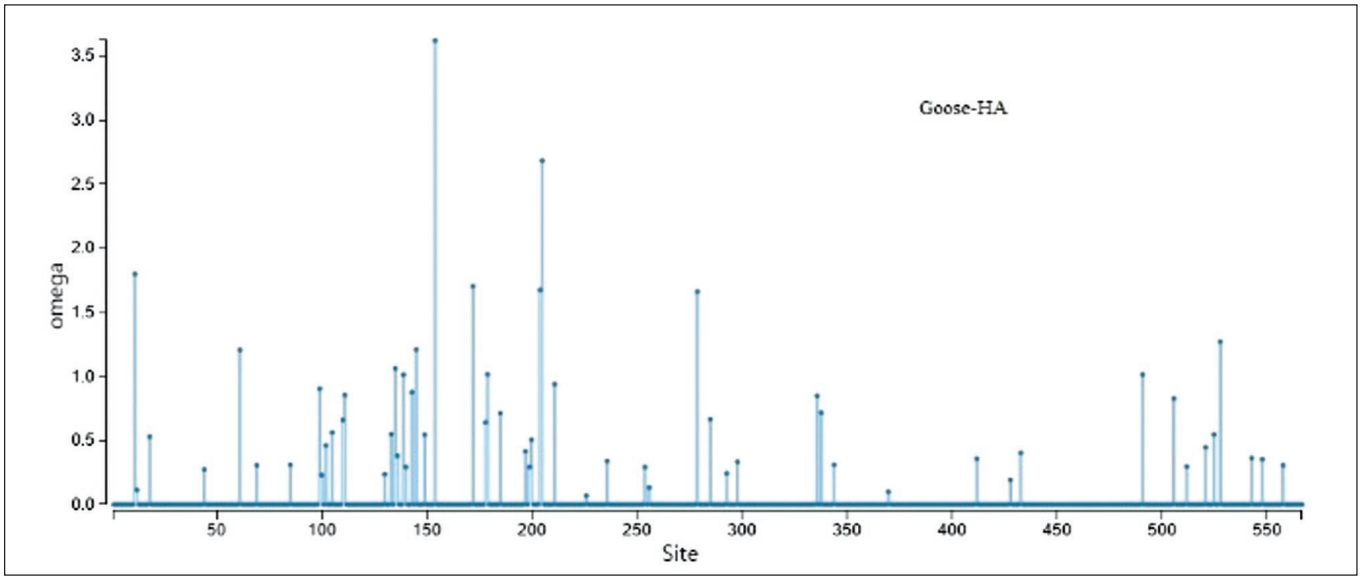

Fig. 4. The positive substitutions per site based on a set of codon aligned nucleotide sequences of chicken H5N1 virus. Cumulated dN-dS are displayed for the codon sites from 1 to 442.

prevalent in the hosts and have advantageous mutations (Suchard *et al.* 2018). In this study, positive selection that had significant signaling of HA gene were detected with 8 positions (87,102,137,145,154,156,157,171) in chicken, 6 positions (156,170,171,172,204,205) in duck and 5 positions (87,131,157,156,171) in goose. There was positive selection in NA gene with 5 positions (16,17,38,60,235) in chicken, 5 positions (44,46,48,56,64) in duck, 3 positions (44,46,369) in goose.

The frequencies of positive selection in the datasets were detected using the SLAC and FEL (Suchard *et al.* 2018). Among the 8 positive selection signals of HA gene in chicken, 3 of them were already predicted by the other 5 selection signals detected in this study have their own significance as evidenced by selection signal sites at 87th position in the epitope region, 102nd and 145th found in antigenic sites, 157th in epitope D and 171st in epitope B

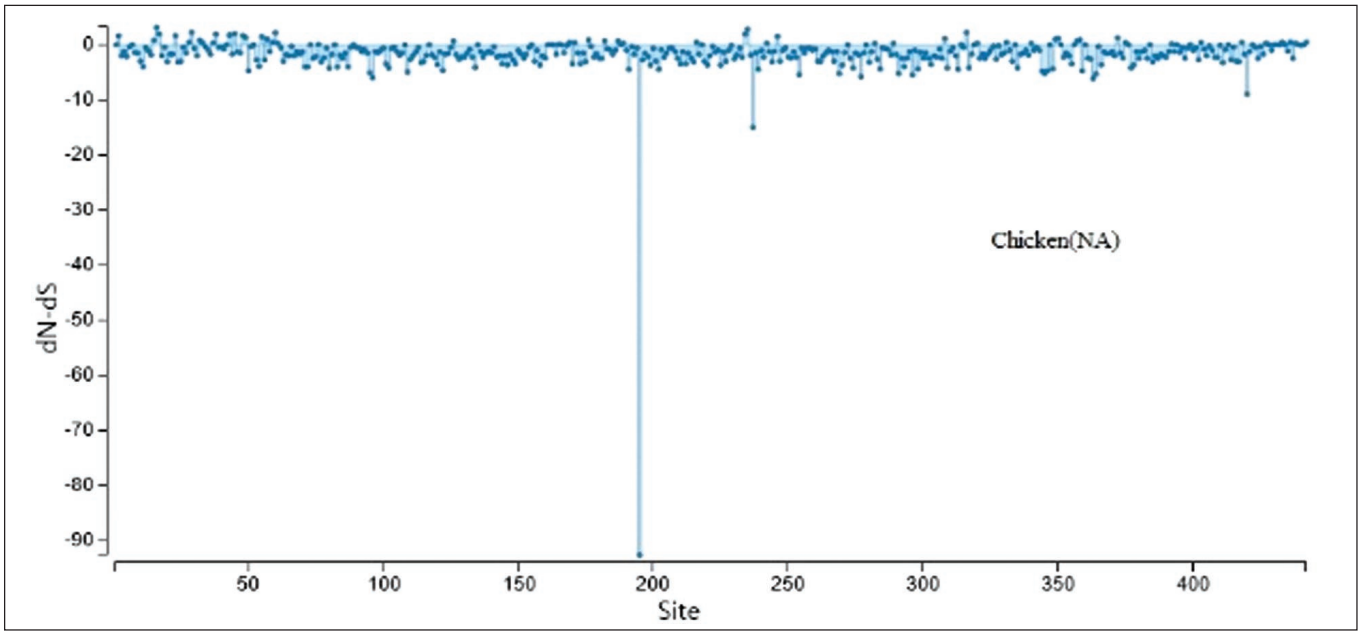

Fig. 5. The positive substitutions per site based on a set of codon aligned nucleotide sequences of duck H5N1 virus. Cumulated dN-dS are displayed for the codon sites from 1 to 449.

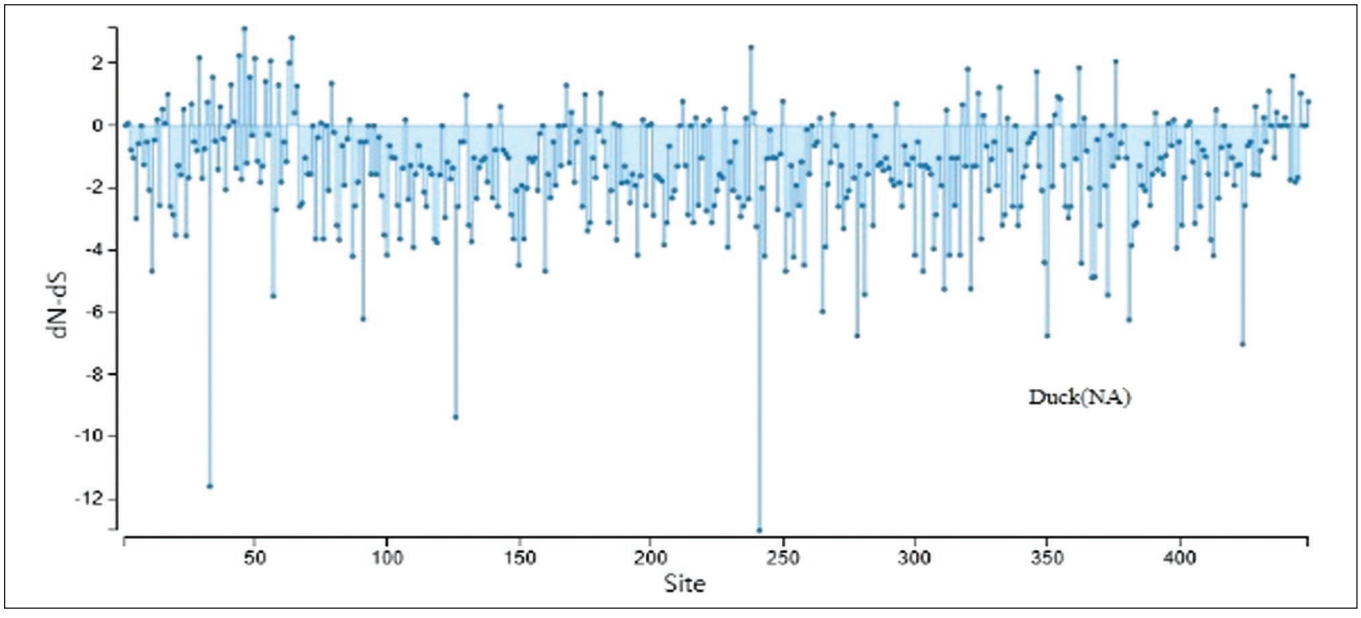

Fig. 6. The positive substitutions per site based on a set of codon aligned nucleotide sequences of goose H5N1 virus. Cumulated omega (dN-dS) is displayed for the codon sites from 1 to 567.

regions (Duvvuri *et al.* 2009). In HA gene of duck avian influenza virus, positive selection sites were found at 170,172, 204 and 205 codon position in this study.

The evolution of H5 and N1 from three avian species revealed that a positive selection played a very critical role, and purifying (negative) selection very largely provided to its viral pathogenicity thereby infecting the hosts. The significant signals show mutations in the HA and NA genes had positive selection and the evolution analyses were convergent, showing that these sites might be evident for adaptation to the host species that were newly evolved (Arafa *et al.* 2016). Positive selection of some codon sites and variants have already been reported previously, variants in the antigenic sites reveal a very captious role and also exhibits that the virus evolves with a high rapidity across the specific geographical region.

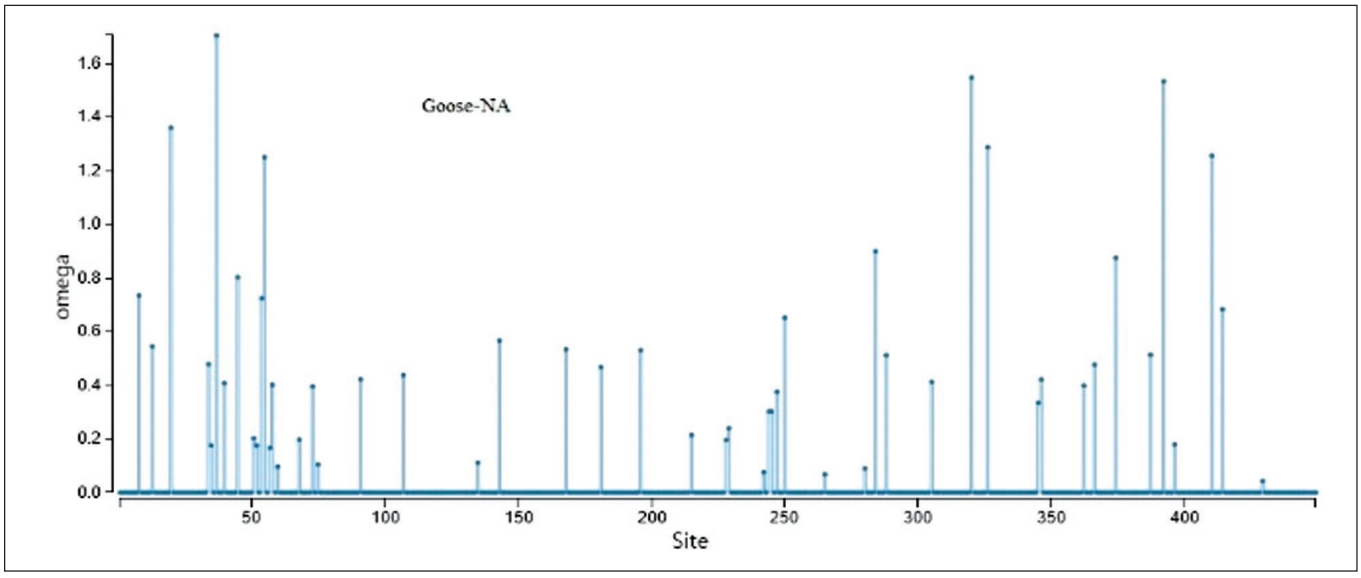

Fig. 7. The positive substitutions per site based on a set of codon aligned nucleotide sequences of goose H5N1 virus. Cumulated omega (dN-dS) is displayed for the codon sites from 1 to 449.

CONCLUSION

To conclude, the evolutionary analysis of the AIV with type A and subtype H5N1 pertaining to the two genes HA and NA from the three avian species (chicken, duck and goose) across Asia revealed that they play very critical role in the evolutionary process. The evolutionary rates and tMRCA observed in this study show heterogeneity among the viral sub-types and also the assessments of nucleotide substitution rate accessible here reveal a more composite picture of the evolutionary progressions driving the evolving avian influenza virus than once thought. Furthermost remarkably, it is strong that a single substitution rate cannot function equally, as the evolutionary rates are consistently lower in the former and which may reflect a greater role for environmental transmission. With the changing environmental conditions, the positive selection signals of these genes would help in interpreting how these genes are mutated and adjusted to the environment. Thus, the analysis of the study was helpful to understand the evolution of H5N1.

REFERENCES

Arafa A, El-Masry I, Kholosy S, Hassan MK, Dauphin G *et al.* (2016) Phylodynamics of avian influenza clade 2.2.1 H5N1 viruses in Egypt. Virol J 13: 49.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17: 540-552.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

Duvvuri VR, Duvvuri B, Cuff WR, Wu GE, Wu J (2009) Role of positive selection pressure on the evolution of H5N1 hemagglutinin. Genomics Proteomics Bioinformatics 7: 47-56.

El-Shesheny R, Bagato O, Kandeil A, Mostafa A, Mahmoud SH *et al.* (2016) Re-emergence of amantadine-resistant variants among highly pathogenic avian influenza H5N1 viruses in Egypt. Infect Genet Evol 46: 102-109.

Fourment M, Holmes EC (2015) Avian influenza virus exhibits distinct evolutionary dynamics in wild birds and poultry. BMC Evol Biol 15: 120.

Gibbs MJ, Armstrong JS, Gibbs AJ (2000) Sister-scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16: 573-582.

Hassan KE, Saad N, Abozeid HH, Shany S, El-Kady MF *et al.* (2020) Genotyping and reassortment analysis of highly pathogenic avian influenza viruses H5N8 and H5N2 from Egypt reveals successive annual replacement of genotypes. Infect Genet Evol 84: 104375.

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinfo 20: 1160-1166.

Klenk HD, Matrosovich M, Stech J (2008) Avian influenza: molecular mechanisms of pathogenesis and host range. Ani Vir Mol Biol 253-301.

Lam HM, Ratmann O, Boni MF (2018) Improved algorithmic complexity for the 3SEQ recombination detection algorithm. Mol Biol Evol 35: 247-251.

Lee DH, Fusaro A, Song CS, Suarez DL, Swayne DE (2016) Poultry vaccination directed evolution of H9N2 low pathogenicity avian influenza viruses in Korea. Virol 488: 225-231.

Li W, Shi W, Qiao H, Ho SY, Luo A *et al.* (2011) Positive selection on hemagglutinin and neuraminidase genes of H1N1 influenza viruses. Virol J 8: 183.

Mahardika GN, Jonas M, Murwijati T, Fitria N, Suartha IN *et al.* (2016) Molecular analysis of hemagglutinin-1 fragment of avian influenza H5N1 viruses isolated from chicken farms in Indonesia from 2008 to 2010. Vet Microbiol 186: 52-58.

Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res Hum Retroviruses 21: 98-102.

Martin D, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. Bioinfo 16: 562-563.

Nonthabenjawan N, Cardona C, Amonsin A, Sreevatsan S (2016) Time-space analysis of highly pathogenic avian influenza H5N2 outbreak in the US. Virol J 13: 147.

Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. Virol 265: 218-225.

Poen MJ, Venkatesh D, Bestebroer TM, Vuong O, Scheuer RD *et al.* (2019) Co-circulation of genetically distinct highly pathogenic avian influenza A clade 2.3.4.4 (H5N6) viruses in wild waterfowl and poultry in Europe and East Asia, 2017-18. Virus Evol 5: 004.

Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinfo 21: 2531-2533.

Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc Natl Acad Sci USA 98: 13757-13762.

Pushko P, Tretyakova I, Hidajat R, Zsak A, Chrzastek K *et al*. (2017) Virus-like particles displaying H5, H7, H9 hemagglutinins and N1 neuraminidase elicit protective immunity to heterologous avian influenza viruses in chickens. Virol 501: 176-182.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in bayesian phylogenetics using tracer 1.7. Syst Biol 67: 901-904.

Rambaut A, Lam TT, Max Carvalho L, Pybus OG (2016) Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2: 007.

Smith JM (1992) Analyzing the mosaic structure of genes. J Mol Evol 34: 126-129.

Suartha IN, Suartini GA, Wirata IW, Dewi NM, Putra GN *et al*. (2018) Intranasal administration of inactivated avian influenza virus of H5N1 subtype vaccine-induced systemic immune response in chicken and mice. Vet World 11: 221-226.

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ *et al.* (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4: 016.

Tian H, Zhou S, Dong L, Van Boeckel TP, Cui Y *et al.* (2015) Correction for Tian *et al.* Avian influenza H5N1 viral and bird migration networks in Asia. Proc Natl Acad Sci USA 112: E2980.

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV *et al.* (2018) Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Mol Biol Evol 35: 773-777.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. Microbiol Rev 56: 152-179.

Zhang Y, Aevermann BD, Anderson TK, Burke DF, Dauphin G *et al*. (2017) Influenza research database: an integrated bioinformatics resource for influenza virus research. Nucleic Acids Res 45: D466-474.

Cite this article as: Suresh KP, Patil S, Indrabalan UB, Sridevi R, Krishnamoorthy P, Rajamani S, Roy P (2020) Evolutionary analysis and detection of positive selection of Hemagglutinin and Neuraminidase genes of H5N1 Avian Influenza virus from chicken, duck and goose across Asia. Explor Anim Med Res 10(2): 169-178.