Infection, Genetics and Evolution 69 (2019) 93–106

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Analysis of codon usage pattern in the viral proteins of chicken anaemia virus and its possible biological relevance

Urmil Dave, Ariktha Srivathsan, Sachin Kumar*

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India

ARTICLE INFO

Keywords:
Chicken anaemia virus
Codon usage
Vaccines
Viral proteins
Pathogenicity

ABSTRACT

Chicken anaemia virus (CAV) is an important poultry pathogen. CAV infection can lead to severe immunosuppression and mortality in birds. The viral proteins (VP1, VP2 and VP3) are major protective immunogens that determine the pathotype of CAV strains. The factors influencing the synonymous codon usage bias and the nucleotide composition of the viral protein genes of CAV have not been studied. In the present study, we have analysed the synonymous codon usage pattern in VP1, VP2 and VP3 genes of CAV. Our results showed that all the genes have a low codon usage bias; however, this bias was slightly higher in the VP2 gene as compared to VP1 and VP3. The major contributing factor for this pattern of codon usage bias in CAV is mutational pressure followed by the host factors. The conclusion was drawn based on codon usage, correspondence analysis, ENC-GC_3 plot and correlation analyses among different indices. The study will help us to understand the codon usage bias of CAV and related single-stranded DNA viruses which could further be used to explore their biology.

1. Introduction

Chicken anaemia virus (CAV) has been an important pathogen for over 45 years owing to its pathogenicity in commercially produced chicken of all ages across the globe. The CAV-induced disease is known to spread by the oral and respiratory routes leading to sub-clinical infection (Rosenberger and Cloud, 1998). Clinical signs such as generalised lymphoid atrophy, increased mortality and severe anaemia in young chicks have been reported. In case of a severe form, the disease could pass on through the progeny chicks by vertical transmission (Adair, 2000). CAV has been identified in most chicken producing regions of the world, including Eastern and Southern Asia (Bhatt et al., 2011; Gholami-Ahangaran, 2015; Kim et al., 2010a; Kye et al., 2013; Zhang et al., 2013a), Africa (AboElkhair et al., 2014; Smuts, 2014; Snoeck et al., 2012), Europe (Engstrom, 1999; McNulty et al., 1990; Olszewska-Tomczyk et al., 2016), North and South America (Craig et al., 2009; Eregae et al., 2014) and Australia (Firth and Imai, 1990). Effect on the economy of these regions, in terms of profitability and productivity of chicken, has been studied in the case of both clinical (McIlroy et al., 1992) and sub-clinical (McConnell et al., 1993; McNulty et al., 1991) forms of the CAV infection.

CAV is a non-enveloped negative sense single-stranded DNA virus belonging to the genus *Gyrovirus* (Schat, 2009). Earlier classified as part of the family *Circoviridae*, CAV is now in the *Anelloviridae* family (Rosario et al., 2017), which mainly consists of viruses with animal hosts. The genome is circular, covalently linked and 2.3 kb long (Noteborn et al., 1991). The ssDNA genome transcribes a polycistronic mRNA with three overlapping open reading frames (ORFs) which further give rise to three proteins, namely VP1, VP2 and VP3. The VP1 is the only structural protein (449 amino acids) of the virus and is responsible for the formation of the viral capsid. In addition, VP1 binds to the viral genomic DNA, directing it to the host cell nucleus (Koch et al., 1995; Lai et al., 2018). The VP2 (216 amino acids) is a non-structural protein and is responsible for scaffolding. The VP2 is also known to have a dual specificity phosphatase domain (Koch et al., 1995; Peters et al., 2002). Additionally, VP2 down-regulates the cell-mediated immune response of the host cells and is considered to be important for the pathogenicity of CAV (Peters et al., 2006). Neutralizing antibodies are induced against both, VP1 and VP2 (Koch et al., 1995). Another 121 amino acid protein, VP3, also known as apoptin, is an important viral protein, known to specifically cause apoptosis of tumour cells (Los et al., 2009; Noteborn et al., 1994; Noteborn, 2004).

It is reported that, in viruses, some codons are preferred over others for the same amino acids (Grantham et al., 1980b; Shackelton et al., 2006). The above phenomenon, pertaining to the preference of one codon over another, is called synonymous codon usage bias. Previous

* Corresponding author.
E-mail address: sachinku@iitg.ac.in (S. Kumar).

https://doi.org/10.1016/j.meegid.2019.01.002
Received 5 July 2018; Received in revised form 19 December 2018; Accepted 2 January 2019
Available online 19 January 2019

studies on synonymous codon usage indicated that distinct coding strategies were not only followed by different genome types, but also in different genes within the same genome (Grantham et al., 1980a). The pattern of synonymous codon usage differentiates the amino acids of varied functional importance. The genes which are highly expressed were found to exhibit higher codon usage bias as compared to the genes which are less expressed (Epstein et al., 2000; Gu et al., 2004). Other poultry viruses such as Newcastle disease virus (Cao et al., 2014; Kumar and Kumar, 2014, 2017), infectious bursal disease virus (Kumar et al., 2015), infectious bronchitis virus (Makhija and Kumar, 2015) and influenza virus (Goni et al., 2012; Li et al., 2010) have been extensively studied for synonymous codon usage for their various genes. However, in the case of CAV, all three proteins have been studied extensively, but the bias in terms of synonymous codons has not been analysed.

2. Materials and methods

2.1. Sequence data and nucleotide composition analysis

The complete genome sequences of 105 CAV isolates, together with their complete GenBank records, were extracted from the GenBank database. The composition of each nucleotide at the third position (A_3%, C_3%, G_3%, T_3%), GC content at the third position (GC_3%), overall GC content (GC%) and dinucleotide frequency of a synonymous codon were calculated using CodonW 1.4.4. (Peden, 1999).

2.2. Recombination and phylogenetic analyses

Potential recombinant events in the CAV sequences were determined with the Recombination Detection Program 4 (RDP 4) Beta software suite (version 4.95) (Martin et al., 2010). The instruction manual of RDP4 was referred to determine the settings. A window length equal to 10% the sequence length was set for analysis. A phylogenetic tree was constructed with a bootstrap of 1000 replicates by the neighbor-joining method, based on the ClustalW alignment (Thompson et al., 1994) using the MEGA5 software (Tamura et al., 2011) as was described in previous studies (Zhang et al., 2013b).

2.2.1. Codon usage analysis

The Relative Synonymous Codon Usage (RSCU) values were calculated as the number of times a codon appears relative to the number of times it would appear, had there been no bias among the codons for a given amino acid. RSCU values are used to determine the characteristics of synonymous codon usage without the confounding influence of amino acid composition and the size of the coding sequence of different gene samples. The synonymous codons with RSCU value > 1.0 have positive codon usage bias and those with RSCU value < 1.0 have a negative codon usage bias. RSCU value equal to 1.0 indicates that codons are chosen randomly and that there is no bias (Sharp and Li, 1986). Moreover, the codons with RSCU value > 1.6 were treated as over-represented while those with RSCU value < 0.6 were treated as under-represented (Wong et al., 2010). RSCU values were calculated using CodonW 1.4.4. Codons AUG and UGG, the only codons for Methionine and Tryptophan respectively, and the termination codons UAA, UAG and UGA are not expected to exhibit any bias and were thus excluded from the analysis.

2.2.2. Codon adaptation index (CAI) analysis

The CAI index is defined as the geometric mean of relative adaptiveness value, which is the ratio of the frequency of use of a given codon to the frequency of use of the optimal codon for the amino acid (Sharp and Li, 1987). CAI values range from 0 to 1, with higher values indicating a higher proportion of the more abundant codons (Sharp and Li, 1986). The high value of CAI refers to higher codon usage bias and expression level (Coghlan and Wolfe, 2000; Duret and Mouchiroud, 1999). CAI calculation requires a reference set of highly expressed known genes and is done using the CAIcal server (Puigbo et al., 2008). The synonymous codon usage pattern of the viral host (*Gallus gallus*) was used as the reference. The said reference dataset for *Gallus gallus* was obtained from the codon usage database (Nakamura et al., 2000).

2.2.3. Effective number of codons (ENC) analysis

ENC is a measure defined as the number of codons that, in case of no codon bias, would give the observed level of codon usage. ENC ranges from 20 to 61 and is used to quantify the absolute codon usage bias of a gene, regardless of its length and the number of amino acids (Hassan et al., 2009). A value of 20 signifies extreme codon usage bias as only one of the possible synonymous codons for the amino acid is used. On the other hand, a value of 61 signifies no bias as all possible synonymous codons are used equally for the amino acid. The greater the codon bias, the smaller the ENC value and significant codon bias are represented by an ENC value less than or equal to 35 (Comeron and Aguade, 1998). ENC values were calculated using CodonW 1.4.4 to evaluate the degree of codon usage bias of CAV genes. Expected ENC values were calculated using the formula: $ENC^{expected} = 2 + s + \left\{\frac{29}{s^2 + (1-s)^2}\right\}$ where s represents GC_3 content (Wright, 1990). If mutation bias solely constrains the codon choice, the genes will lie on or just below the curve of the expected ENC values.

2.2.4. Correspondence analysis (CoA) of codon usage

Correspondence Analysis (CoA) is a multivariate statistical method instrumental in exploring the relationship between variables and samples, which is also employed to gain insight into codon usage patterns (Grantham et al., 1980a). CodonW 1.4.4 was used to perform CoA on RSCU values to analyse the major trends in codon usage patterns of CAV. RSCU values of 59 codons (excluding Met, Trp and the three stop codons) corresponded to 59 dimensions and each sequence was represented using this 59-dimensional vector.

2.2.5. Relative dinucleotide abundance

The relative dinucleotide abundance, which aids in establishing the relationship of dinucleotide bias with codon usage bias, was calculated using a previously described method (Karlin and Burge, 1995). Expected dinucleotide values were calculated assuming random association of nucleotides from the observed frequencies for every sequence. The odds ratio was calculated using the formula:

$$P_{xy} = f_{xy}/f_y f_x,$$

where f_x denotes the frequency of nucleotide X, f_y denotes the frequency of nucleotide Y, $f_y f_x$ denotes the expected frequency of the dinucleotide XY, and f_{xy} denotes the frequency of the dinucleotide XY calculated using CodonW 1.4.4. A ratio > 1.23 signifies over-representation and a ratio < 0.78 signifies under-representation in terms of relative abundance compared with a random association of mononucleotides.

2.2.6. Grand average of hydropathy (GRAVY)

GRAVY calculates the hydropathic character of a protein by analysing the arithmetic mean of the sum of the hydrophobic indices of each amino acid (Kyte and Doolittle, 1982). The Positive GRAVY value indicates that the protein is polar, while a negative value indicates that the protein is non-polar. GRAVY values were calculated using CodonW 1.4.4.

2.2.7. Correlation analysis

Correlation analysis was performed using Pearson correlation to identify the relation between nucleotide composition, two axes of CoA, and variables like GRAVY, Aromaticity (AROMO), ENC and GC3.

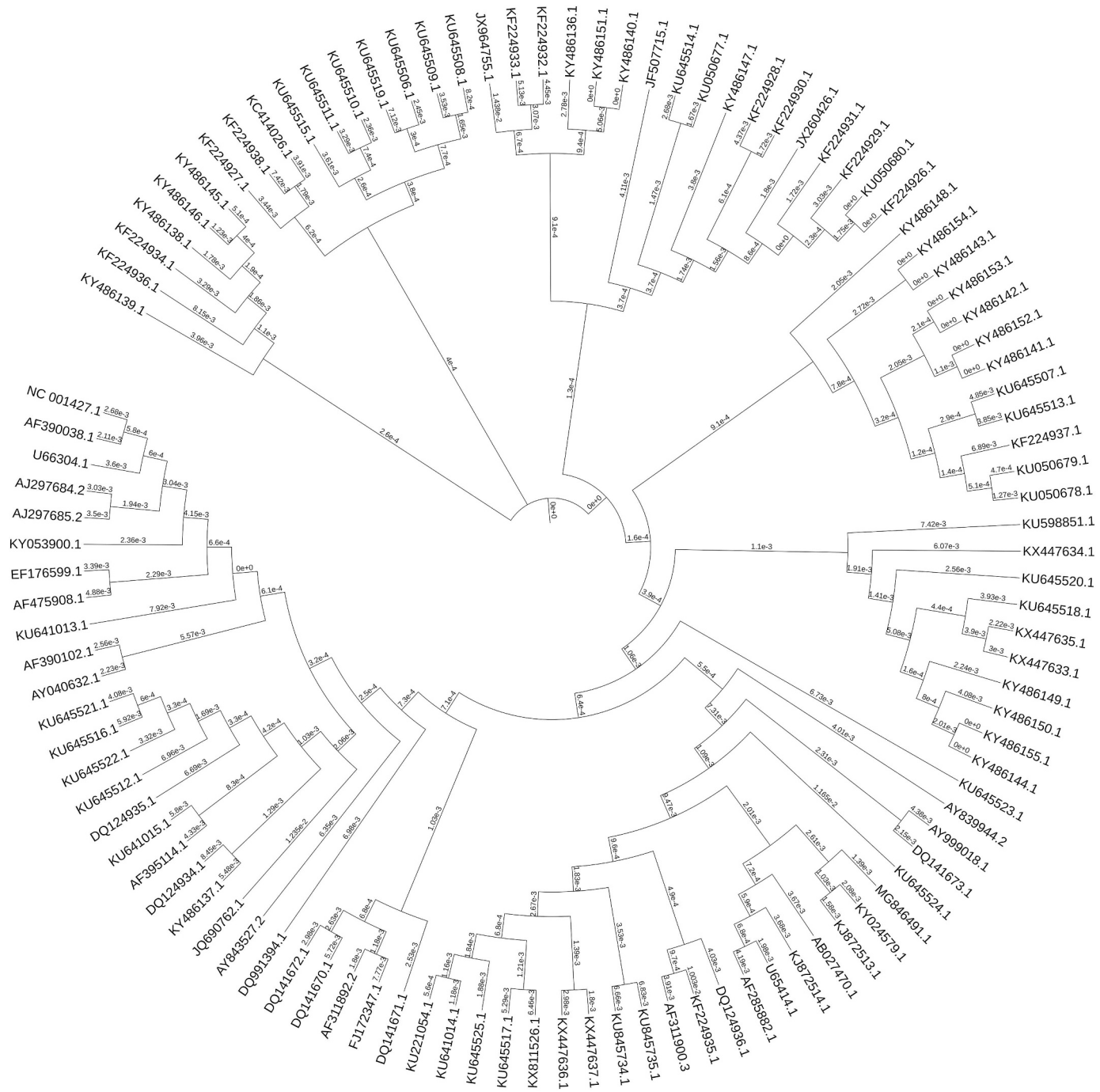

Fig. 1. A phylogenetic tree of 105 CAV sequences was constructed with a bootstrap of 1000 replicates by the neighbor-joining method, based on the Clustal W alignment.

3. Results

3.1. Recombination and phylogenetic analyses

The analysis of 105 CAV sequences by RDP4 Beta software showed no recombination event. The phylogenetic analysis of all the sequences performed using MEGA5 software showed alignment with different clades with varied bootstrap values (Fig. 1).

3.2. Nucleotide composition analysis

Overall nucleotide composition of coding sequences greatly influences preference for one type of codon over another (Jenkins and Holmes, 2003). On the analysis of 105 complete genome sequences of CAV by considering ORFs for VP1, VP2 and VP3 separately, the mean GC% for all the three genes was fairly alike with a mean of 55.32%, 55.81%, and 55.32%, respectively. In all of the genes, $\%G_3$ was the highest and $\%T_3$ was the lowest. It was observed that in VP1, C/G-ended codons were greatly preferred, while they were only slightly preferred over A/U-ended codons in VP2 and VP3 (Table 1).

3.3. Analysis of relative synonymous codon usage (RSCU) characteristics

RSCU analysis was performed to assess the patterns of synonymous codon usage in the viral proteins of CAV, and to decrypt the extent to which C/G-ended codons (for VP1) and C/A-ended codons (for VP2 and

Table 1
Details of VP1 (A), VP2 (B), and VP3 (C) genes used for analysis of synonymous codon usage.

(A)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
1	U65414.1	NA	propagated in MDCC-MSB1 cells; replicative form of the virus was isolated and cloned into a bacterial plasmid		55.2	58.2	55.58	21.29	38.1	29.75	33.55	0.753
2	AF311892.2	NA	*Gallus gallus*	NA	55.5	59.1	53.5	20.79	39.33	29.23	33.23	0.773
3	AF311900.3	NA	Gallus gallus	NA	55.4	58.9	54.44	20.79	39.04	29.36	33.44	0.762
4	AY040632.1	NA	Gallus gallus	Malaysia	55.5	59.7	53.13	21.13	39.15	28.09	34.29	0.78
5	AF285882.1	NA	Gallus gallus	Malaysia	54.9	57.5	56.57	21.85	37.82	30.06	32.91	0.75
6	AF475908.1	NA	Gallus gallus	China: Harbin	55.2	58.4	53.99	20.79	39.33	30.15	32.27	0.772
7	AF390038.1	NA	Gallus gallus	Malaysia	55.6	60	53.6	20.85	39.72	28.09	34.29	0.778
8	AF395114.1	NA	Gallus gallus	NA	55.7	59.9	53.46	19.61	39.5	29.45	34.19	0.782
9	AF390102.1	NA	Gallus gallus	Malaysia	55.2	58.5	53.85	21.69	38.87	29.01	32.91	0.781
10	AY839944.2	NA	NA	China: Hebei Province	55.2	58.6	55.58	21.91	39.04	28.62	32.91	0.762
11	AY843527.2	NA	Gallus gallus breed Silkies	China: Tianjin	55.8	59.8	53.17	20.51	40.17	28.62	33.12	0.774
12	AY999018.1	NA	NA	China: Shandong Province	55.3	58.8	55.24	20.79	39.89	29.54	32.17	0.762
13	DQ124934.1	NA	NA	China:Beijing	55.8	59.8	53.46	19.83	39.39	29.36	33.86	0.778
14	DQ124935.1	NA	NA	China:Anhui Province	55.5	59.4	54.68	21.35	39.04	28.13	34.08	0.776
15	DQ124936.1	NA	NA	China:Anhui Province	55.6	59.4	54.31	21.57	38.94	27.83	34.08	0.76
16	DQ141670.1	NA	NA	China: Shanghai	55.5	59.7	53.8	20.45	40.34	28.83	32.91	0.777
17	DQ141671.1	NA	NA	China: Shanghai	55.5	59.3	54.55	20.22	40.17	29.54	32.48	0.767
18	DQ141672.1	NA	NA	China: Henan Province	55	57.9	55.42	20.51	39.33	31.08	31.53	0.766
19	DQ141673.1	NA	NA	China: Shandong Pronvince	55.2	58.6	55.17	21.07	39.61	29.54	32.17	0.765
20	DQ991394.1	NA	NA	USA	55.5	59.1	53.29	20.22	40.45	29.85	31.95	0.779
21	EF176599.1	NA	NA	China	55.2	58.4	54.1	20.79	39.04	30.15	32.59	0.769
22	FJ172347.1	NA	broiler chicken	China	55.7	60	53.34	20.45	39.5	28.31	34.19	0.774
23	JF507715.1	NA	Gallus gallus	South Korea	55.2	58.6	56.8	20.79	39.33	29.94	32.69	0.765
24	JQ690762.1	Feces	*Homo sapiens*	China	55.2	58.6	53.6	21.35	39.04	29.32	33.12	0.78
25	JX260426.1	NA	Gallus gallus	China	55.3	58.5	57.41	21.41	38.87	29.32	33.01	0.757
26	JX964755.1	NA	Gallus gallus	China	55.4	59.1	54.78	20.51	38.76	29.54	33.87	0.767
27	KC414026.1	NA	Cat	China	55.8	59.8	55.28	20.11	39.66	28.92	33.55	0.763
28	KF224926.1	NA	Gallus gallus	China: Guangzhou	55.2	58.5	56.03	21.41	38.87	29.32	33.01	0.759
29	KF224927.1	NA	Gallus gallus	China: Guangzhou	55.5	59.3	55.3	20.51	39.61	29.23	33.23	0.76
30	KF224928.1	NA	Gallus gallus	China: Guangzhou	55	58.4	57.42	21.85	38.94	29.1	32.8	0.759
31	KF224929.1	NA	Gallus gallus	China: Guangzhou	55.1	58.3	55.8	21.41	38.87	29.63	32.69	0.759
32	KF224930.1	NA	Gallus gallus	China: Guangzhou	55.2	58.7	56.46	21.41	38.87	29.01	33.33	0.761
33	KF224931.1	NA	Gallus gallus	China: Guangzhou	55.3	58.5	56.09	21.41	38.87	29.32	33.01	0.759
34	KF224932.1	NA	Gallus gallus	China: Guangzhou	55.3	58	56.11	20.67	38.55	30.58	32.38	0.756
35	KF224933.1	NA	Gallus gallus	China: Guangzhou	54.9	57.9	56.1	20.79	38.76	30.77	32.27	0.76
36	KF224934.1	NA	Gallus gallus	China: Guangzhou	55.3	58.8	55.11	20.79	39.33	29.54	32.91	0.763
37	KF224935.1	NA	Gallus gallus	China: Guangzhou	55.3	58.2	54.93	21.29	38.38	29.66	33.12	0.76
38	KF224936.1	NA	Gallus gallus	China: Guangzhou	55.2	58.8	55.68	20.56	39.44	29.85	32.91	0.761
39	KF224937.1	NA	Gallus gallus	China: Guangzhou	55.4	59.1	54.81	21.07	39.04	28.92	33.55	0.761
40	KF224938.1	NA	Gallus gallus	China: Guangzhou	55.3	58.4	56.23	20.45	39.5	30.37	31.85	0.756
41	KJ872513.1	thymus	Gallus gallus; commercial broiler	Argentina	54.9	57.7	55.65	21.57	38.1	30.06	32.91	0.753
42	KJ872514.1	thymus	Gallus gallus; commercial broiler	Argentina	55.1	58.5	55.93	21.79	37.99	28.92	33.97	0.758
43	KU845734.1	NA	Gallus gallus	China	55.4	58.5	55.12	21.97	38.31	28.75	33.76	0.753
44	KU845735.1	NA	Gallus gallus	China	55.2	58.2	54.27	22.07	37.99	28.75	33.55	0.761
45	KU221054.1	NA	Gallus gallus	China	55.2	58.8	54.75	21.57	38.38	28.66	33.97	0.762
46	KU050677.1	NA	Gallus gallus	NA	55.1	58.1	55.49	20.79	39.04	30.46	32.27	0.763
47	KU050678.1	NA	Gallus gallus	NA	55.3	58.8	55.32	20.79	39.33	29.54	32.91	0.763
48	KU050679.1	NA	Gallus gallus	NA	55.2	58.6	55.51	20.51	39.33	30.15	32.59	0.762
49	KU050680.1	NA	Gallus gallus	NA	55.2	58.5	56.03	21.41	38.87	29.32	33.01	0.759
50	KU598851.1	NA	Gallus gallus	China	54.6	56.9	55.98	21.01	38.94	31.89	30.87	0.763
51	KU641013.1	NA	Gallus gallus	China	55.3	59	53.06	21.07	38.76	29.1	33.76	0.778
52	KU641014.1	NA	Gallus gallus	China	55.2	58.8	54.84	21.57	38.1	28.66	34.29	0.761
53	KU641015.1	NA	Gallus gallus	China	55.7	59.8	54.15	20.11	39.39	29.05	34.08	0.777
54	KU645506.1	NA	Gallus gallus	China	55.4	59.1	55.47	20.45	39.22	29.63	33.33	0.765
55	KU645507.1	NA	Gallus gallus	China	55	58.4	55.5	20.79	39.33	30.25	32.37	0.766
56	KU645508.1	NA	Gallus gallus	China	55.5	59.8	55.08	19.89	39.78	29.32	33.65	0.765
57	KU645509.1	NA	Gallus gallus	China	55.5	59.8	55.25	19.94	39.61	29.23	33.87	0.766
58	KU645510.1	NA	Gallus gallus	China	55.5	59.6	55.24	20.79	39.33	28.62	33.87	0.765
59	KU645511.1	NA	Gallus gallus	China	55.7	59.3	55.55	21.07	39.04	28.53	33.76	0.763
60	KU645512.1	NA	Gallus gallus	China	55.6	59.9	53.83	20.11	39.66	28.92	33.97	0.78
61	KU645513.1	NA	Gallus gallus	China	55.4	58.5	55.52	20.85	39.44	29.85	32.27	0.766
62	KU645514.1	NA	Gallus gallus	China	55.2	58.6	55.62	20.95	39.39	29.72	32.48	0.767
63	KU645515.1	NA	Gallus gallus	China	55.5	59.3	54.77	20.51	39.33	29.23	33.55	0.767
64	KU645516.1	NA	Gallus gallus	China	55.4	58.9	54.7	21.23	38.83	28.92	33.65	0.776
65	KU645517.1	NA	Gallus gallus	China	55.3	58.7	54.62	21.57	38.66	28.83	33.55	0.763

(*continued on next page*)

Table 1 (*continued*)

(A)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
66	KU645518.1	NA	Gallus gallus	China	55.2	58.4	55.4	21.07	38.76	29.75	32.8	0.772
67	KU645519.1	NA	Gallus gallus	China	55.3	59.4	55.7	20.39	39.39	29.23	33.65	0.769
68	KU645520.1	NA	Gallus gallus	China	55.8	60	53.91	20.79	39.04	28	34.82	0.774
69	KU645521.1	NA	Gallus gallus	China	55.3	58.9	54.71	21.01	38.38	29.14	34.19	0.773
70	KU645522.1	NA	Gallus gallus	China	55.2	58.5	55.2	21.29	38.1	29.45	33.76	0.774
71	KU645523.1	NA	Gallus gallus	China	55.6	59.8	54.15	20.45	39.22	28.7	34.19	0.764
72	KU645524.1	NA	Gallus gallus	China	55.2	58.3	54.43	21.07	39.89	29.85	31.21	0.764
73	KU645525.1	NA	Gallus gallus	China	55.4	58.7	55.24	21.63	38.48	28.66	33.65	0.761
74	KY024579.1	feces	Gallus gallus	Brazil	54.9	58	55.5	21.57	38.38	29.75	32.91	0.753
75	KX811526.1	NA	Gallus gallus	China	55.4	58.7	54.54	21.57	38.66	28.66	33.33	0.763
76	KX447633.1	liver	Gallus gallus	China	55.2	58.6	55.61	21.57	38.66	29.01	33.33	0.772
77	KX447634.1	liver	Gallus gallus	China	55.5	59.4	55.04	21.13	39.44	28.57	33.44	0.771
78	KX447635.1	liver	Gallus gallus	China	55.2	58.6	55.38	21.35	38.76	29.23	33.23	0.77
79	KX447636.1	liver	Gallus gallus	China	55.3	58.7	54.62	21.57	38.66	28.83	33.55	0.763
80	KX447637.1	liver	Gallus gallus	China	55	58.3	54.8	22.13	38.1	28.66	33.65	0.761
81	KY053900.1	NA	Gallus gallus	India	55.9	60.5	52.58	20.51	40.73	27.69	33.55	0.783
82	KY486136.1	liver; spleen	Gallus gallus	China	55.5	59.3	55.21	20.51	39.33	29.23	33.55	0.764
83	KY486137.1	liver; spleen	Gallus gallus	China	55.2	58.2	54.17	20.45	38.1	30.67	33.55	0.77
84	KY486138.1	liver; spleen	Gallus gallus	China	55.3	59.1	55.29	21.07	39.33	28.92	33.23	0.764
85	KY486139.1	liver; spleen	Gallus gallus	China	55.4	59.2	55.63	20.85	39.44	29.01	33.33	0.761
86	KY486140.1	liver; spleen	Gallus gallus	China	55.3	58.8	55.47	20.51	39.61	29.85	32.59	0.765
87	KY486141.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.33	20.51	39.61	29.54	32.91	0.764
88	KY486142.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.33	20.51	39.61	29.54	32.91	0.764
89	KY486143.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.04	20.51	39.33	29.54	33.23	0.767
90	KY486144.1	liver; spleen	Gallus gallus	China	55.2	58.6	55.39	21.57	38.66	29.01	33.33	0.772
91	KY486145.1	liver; spleen	Gallus gallus	China	55.3	59.1	55.47	20.51	39.33	29.54	33.23	0.767
92	KY486146.1	liver; spleen	Gallus gallus	China	55.4	59.3	55.17	20.51	39.33	29.23	33.55	0.767
93	KY486147.1	liver; spleen	Gallus gallus	China	55.2	58.5	55.81	20.85	39.44	29.94	32.37	0.761
94	KY486148.1	liver; spleen	Gallus gallus	China	55.2	58.4	55.22	21.07	39.04	29.85	32.59	0.762
95	KY486149.1	liver; spleen	Gallus gallus	China	55.1	58.1	55.55	21.85	38.66	29.32	32.69	0.772
96	KY486150.1	liver; spleen	Gallus gallus	China	55.1	58.1	55.56	21.85	38.38	29.32	33.01	0.773
97	KY486151.1	liver; spleen	Gallus gallus	China	55.3	58.8	55.47	20.51	39.61	29.85	32.59	0.765
98	KY486152.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.33	20.51	39.61	29.54	32.91	0.764
99	KY486153.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.33	20.51	39.61	29.54	32.91	0.764
100	KY486154.1	liver; spleen	Gallus gallus	China	55.4	59.1	55.04	20.51	39.33	29.54	33.23	0.767
101	KY486155.1	liver; spleen	Gallus gallus	China	55.2	58.6	55.39	21.57	38.66	29.01	33.33	0.772
102	MG846491.1	feces	Gallus gallus	Brazil	54.8	57.2	55.76	21.57	38.1	30.67	32.27	0.754
103	AB027470.1	NA	wild type	NA	55.1	58.5	55.25	21.29	38.38	29.45	33.65	0.753
104	AJ297684.2	NA	Virus cloned from a Cux-1 virus pool that has recieved 310 passages in cell culture	Germany:Cuxhaven	55.8	60.3	53.21	20.39	39.11	28	34.82	0.78
105	AJ297685.2	NA	Virus cloned from a Cux-1 virus pool that has recieved 310 passages in cell culture	Germany:Cuxhaven	55.9	60.3	53.19	20.39	39.39	28	34.5	0.781
	MEAN				55.3	58.8	55.01	20.95	39.09	29.32	33.20	0.766

(*continued on next page*)

Table 1 (*continued*)

(B)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
1	U65414.1	NA	propagated in MDCC-MSB1 cells; replicative form of the virus was isolated and cloned into a bacterial plasmid	NA	55.7	57.1	52.19	21.14	43.43	32.32	28.39	0.777
2	AF311892.2	NA	Gallus gallus	NA	55.9	57.6	51.81	21.14	43.43	31.71	29.03	0.777
3	AF311900.3	NA	Gallus gallus	NA	55.9	57.6	51.62	21.14	44	31.71	28.39	0.778
4	AY040632.1	NA	Gallus gallus	Malaysia	56	57.1	52.24	21.02	43.18	32.32	28.39	0.773
5	AF285882.1	NA	Gallus gallus	Malaysia	55.7	57.1	51.92	21.14	43.43	32.32	28.39	0.777
6	AF475908.1	NA	Gallus gallus	China: Harbin	55.9	57.6	51.83	21.14	44	31.71	28.39	0.777
7	AF390038.1	NA	Gallus gallus	Malaysia	55.7	57.1	51.16	21.14	44	32.32	27.74	0.778
8	AF395114.1	NA	Gallus gallus	NA	55.7	57.1	51.9	21.14	44	32.32	27.74	0.777
9	AF390102.1	NA	Gallus gallus	Malaysia	56	57.6	51.3	21.14	44	31.71	28.39	0.774
10	AY839944.2	NA	NA	China: Hebei Province	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
11	AY843527.2	NA	Gallus gallus breed Silkies	China: Tianjin	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
12	AY999018.1	NA	NA	China: Shandong Province	55.7	57.6	51.24	21.14	44	31.71	28.39	0.778
13	DQ124934.1	NA	NA	China:Beijing	55.6	56.7	52.25	21.14	43.43	32.93	27.74	0.774
14	DQ124935.1	NA	NA	China:Anhui Province	56.2	58.6	51.15	20	45.14	31.71	28.39	0.78
15	DQ124936.1	NA	NA	China:Anhui Province	56	58.1	51.41	21.14	44	31.1	29.03	0.779
16	DQ141670.1	NA	NA	China: Shanghai	55.6	57.1	52.25	21.14	43.43	32.32	28.39	0.778
17	DQ141671.1	NA	NA	China: Shanghai	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
18	DQ141672.1	NA	NA	China: Henan Province	55.7	57.6	51.24	21.14	44	31.71	28.39	0.778
19	DQ141673.1	NA	NA	China: Shandong Pronvince	55.6	57.1	51.03	21.14	44	32.32	27.74	0.778
20	DQ991394.1	NA	NA	USA	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
21	EF176599.1	NA	NA	China	56	57.6	51.33	21.02	43.75	31.71	28.39	0.777
22	FJ172347.1	NA	broiler chicken	China	55.4	57.1	53.23	21.14	43.43	32.32	28.39	0.775
23	JF507715.1	NA	Gallus gallus	South Korea	55.7	57.1	52.13	21.71	43.43	31.71	28.39	0.777
24	JQ690762.1	Feces	Homo sapiens	China	56.2	58.1	51.73	20.45	44.89	31.71	27.74	0.78
25	JX260426.1	NA	Gallus gallus	China	56	58.1	51.07	21.14	44	31.1	29.03	0.775
26	JX964755.1	NA	Gallus gallus	China	55.7	57.1	51.44	21.14	44	32.32	27.74	0.778
27	KC414026.1	NA	Cat	China	56	57.6	51.61	21.14	44	31.71	28.39	0.776
28	KF224926.1	NA	Gallus gallus	China: Guangzhou	55.7	57.4	51.36	21.26	44.25	32.1	28.1	0.778
29	KF224927.1	NA	Gallus gallus	China: Guangzhou	55.7	57.6	51.37	21.14	44.57	31.71	27.74	0.777
30	KF224928.1	NA	Gallus gallus	China: Guangzhou	56	58.1	50.26	20.45	43.75	31.52	28.85	0.78
31	KF224929.1	NA	Gallus gallus	China: Guangzhou	55.7	57.6	51.48	21.26	44.25	31.71	28.39	0.777
32	KF224930.1	NA	Gallus gallus	China: Guangzhou	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
33	KF224931.1	NA	Gallus gallus	China: Guangzhou	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
34	KF224932.1	NA	Gallus gallus	China: Guangzhou	56	57.6	51.24	21.14	44	31.71	28.39	0.775
35	KF224933.1	NA	Gallus gallus	China: Guangzhou	56	57.6	51.23	21.02	43.75	31.71	28.39	0.777
36	KF224934.1	NA	Gallus gallus	China: Guangzhou	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
37	KF224935.1	NA	Gallus gallus	China: Guangzhou	55.9	57.1	52.23	21.14	44	32.12	27.56	0.777
38	KF224936.1	NA	Gallus gallus	China: Guangzhou	55.7	57.1	51.92	21.14	43.43	32.32	28.39	0.777
39	KF224937.1	NA	Gallus gallus	China: Guangzhou	55.7	57.1	52.01	21.71	43.43	31.71	28.39	0.781
40	KF224938.1	NA	Gallus gallus	China: Guangzhou	56	58.1	52.6	21.14	44.57	30.91	28.03	0.778
41	KJ872513.1	thymus	Gallus gallus; commercial broiler	Argentina	55.7	57.1	51.92	21.14	43.43	32.32	28.39	0.777
42	KJ872514.1	thymus	Gallus gallus; commercial broiler	Argentina	55.7	57.1	51.92	21.14	43.43	32.32	28.39	0.777
43	KU845734.1	NA	Gallus gallus	China	55.9	58.1	52.47	20.57	44.57	31.71	28.39	0.783
44	KU845735.1	NA	Gallus gallus	China	55.6	57.1	52.69	21.14	44	32.32	27.74	0.781
45	KU221054.1	NA	Gallus gallus	China	56.2	58.6	51.37	20.57	44.57	31.1	29.03	0.785
46	KU050677.1	NA	Gallus gallus	NA	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
47	KU050678.1	NA	Gallus gallus	NA	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
48	KU050679.1	NA	Gallus gallus	NA	55.9	57.6	51.27	21.14	44	31.9	28.57	0.781
49	KU050680.1	NA	Gallus gallus	NA	55.7	57.4	51.36	21.26	44.25	32.1	28.1	0.778
50	KU598851.1	NA	Gallus gallus	China	55.7	57.6	51.75	21.14	44	31.71	28.39	0.778
51	KU641013.1	NA	Gallus gallus	China	55.7	56.7	52.38	21.14	43.43	32.73	27.56	0.774
52	KU641014.1	NA	Gallus gallus	China	56.2	58.6	51.37	20.57	44.57	31.1	29.03	0.785
53	KU641015.1	NA	Gallus gallus	China	55.7	57.1	51.87	21.71	42.86	31.71	29.03	0.776
54	KU645506.1	NA	Gallus gallus	China	55.6	57.1	52.27	21.14	44	32.32	27.74	0.777
55	KU645507.1	NA	Gallus gallus	China	55.9	58.1	51.98	21.14	44	31.1	29.03	0.783
56	KU645508.1	NA	Gallus gallus	China	55.6	56.7	51.18	21.26	44.25	32.73	26.92	0.778
57	KU645509.1	NA	Gallus gallus	China	55.7	56.7	53.09	21.26	44.25	32.53	26.75	0.774
58	KU645510.1	NA	Gallus gallus	China	55.6	57.1	52.35	21.26	44.25	32.32	27.74	0.779
59	KU645511.1	NA	Gallus gallus	China	55.6	56.7	52.14	21.14	44	32.93	27.1	0.779
60	KU645512.1	NA	Gallus gallus	China	55.6	57.6	52.24	21.26	43.68	31.9	29.22	0.777
61	KU645513.1	NA	Gallus gallus	China	56	58.1	51.22	20.57	44.57	31.71	28.39	0.778
62	KU645514.1	NA	Gallus gallus	China	56	58.1	50.91	20.57	44.57	31.71	28.39	0.778

(*continued on next page*)

Table 1 (*continued*)

(B)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
63	KU645515.1	NA	Gallus gallus	China	55.9	57.6	51.75	21.14	44.57	31.71	27.74	0.776
64	KU645516.1	NA	Gallus gallus	China	55.7	57.1	52.44	21.14	43.43	32.32	28.39	0.777
65	KU645517.1	NA	Gallus gallus	China	55.9	58.1	52.47	20.57	44.57	31.71	28.39	0.783
66	KU645518.1	NA	Gallus gallus	China	55.6	57.1	52.32	21.14	43.43	32.32	28.39	0.774
67	KU645519.1	NA	Gallus gallus	China	55.7	57.1	51.28	21.14	44	32.32	27.74	0.777
68	KU645520.1	NA	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
69	KU645521.1	NA	Gallus gallus	China	55.6	57.1	51.92	21.14	43.43	32.52	28.57	0.779
70	KU645522.1	NA	Gallus gallus	China	55.7	57.1	52.39	21.14	43.43	32.32	28.39	0.777
71	KU645523.1	NA	Gallus gallus	China	55.6	56.2	51.34	21.02	43.18	33.54	27.1	0.776
72	KU645524.1	NA	Gallus gallus	China	55.6	57.1	53.03	20.57	44	32.93	27.74	0.777
73	KU645525.1	NA	Gallus gallus	China	56.2	58.6	51.73	20.57	44.57	31.1	29.03	0.784
74	KY024579.1	feces	Gallus gallus	Brazil	55.6	56.7	52.99	21.14	42.86	32.93	28.39	0.777
75	KX811526.1	NA	Gallus gallus	China	55.7	57.6	52.22	20.57	44.57	32.32	27.74	0.779
76	KX447633.1	liver	Gallus gallus	China	55.7	57.1	51.9	21.14	44	32.32	27.74	0.777
77	KX447634.1	liver	Gallus gallus	China	55.6	57.1	51.75	21.14	44	32.32	27.74	0.777
78	KX447635.1	liver	Gallus gallus	China	55.7	57.1	51.75	20.69	44.25	32.73	27.56	0.776
79	KX447636.1	liver	Gallus gallus	China	55.9	57.1	52.26	21.14	44.57	32.12	26.92	0.778
80	KX447637.1	liver	Gallus gallus	China	56	57.6	51.84	21.14	44.57	31.52	27.56	0.779
81	KY053900.1	NA	Gallus gallus	India	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
82	KY486136.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.43	21.14	44.57	31.71	27.74	0.778
83	KY486137.1	liver; spleen	Gallus gallus	China	55.7	57.1	52.39	21.14	43.43	32.32	28.39	0.777
84	KY486138.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
85	KY486139.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
86	KY486140.1	liver; spleen	Gallus gallus	China	56	58.1	51.43	21.14	44	31.1	29.03	0.781
87	KY486141.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
88	KY486142.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
89	KY486143.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.81	21.14	43.43	31.71	29.03	0.777
90	KY486144.1	liver; spleen	Gallus gallus	China	55.7	57.1	52.39	21.71	43.43	31.71	28.39	0.777
91	KY486145.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
92	KY486146.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
93	KY486147.1	liver; spleen	Gallus gallus	China	55.7	57.1	51.9	21.14	44	32.32	27.74	0.777
94	KY486148.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
95	KY486149.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
96	KY486150.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
97	KY486151.1	liver; spleen	Gallus gallus	China	56	58.1	51.43	21.14	44	31.1	29.03	0.781
98	KY486152.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
99	KY486153.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.38	21.14	44	31.71	28.39	0.778
100	KY486154.1	liver; spleen	Gallus gallus	China	55.9	57.6	51.81	21.14	43.43	31.71	29.03	0.777
101	KY486155.1	liver; spleen	Gallus gallus	China	55.7	57.1	52.39	21.71	43.43	31.71	28.39	0.777
102	MG846491.1	feces	Gallus gallus	Brazil	55.7	57.1	51.92	21.14	43.43	32.32	28.39	0.777
103	AB027470.1	NA	wild type	NA	55.6	56.7	52.44	21.14	43.43	32.93	27.74	0.776
104	AJ297684.2	NA	NA	Germany:Cuxhaven	55.7	56.7	52.86	21.59	42.61	32.32	28.39	0.772
105	AJ297685.2	NA	NA	Germany:Cuxhaven	55.7	57.1	52.81	21.02	43.18	32.32	28.39	0.776
	MEAN				55.8	57.4	51.76	21.10	43.92	31.95	28.25	0.778

(*continued on next page*)

Table 1 (*continued*)

(C)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
1	U65414.1	NA	propagated in MDCC-MSB1 cells; replicative form of the virus was isolated and cloned into a bacterial plasmid	NA	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
2	AF311892.2	NA	Gallus gallus	NA	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.779
3	AF311900.3	NA	Gallus gallus	NA	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
4	AY040632.1	NA	Gallus gallus	Malaysia	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.78
5	AF285882.1	NA	Gallus gallus	Malaysia	54.9	51.7	53.7	24.53	32.08	31.37	29.47	0.782
6	AF475908.1	NA	Gallus gallus	China: Harbin	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
7	AF390038.1	NA	Gallus gallus	Malaysia	55.6	51.7	53.7	24.53	32.08	31.37	29.47	0.782
8	AF395114.1	NA	Gallus gallus	NA	55.7	51.7	53.4	24.53	32.08	31.68	29.79	0.782
9	AF390102.1	NA	Gallus gallus	Malaysia	55.2	51.7	54	24.53	32.08	31.07	29.17	0.78
10	AY839944.2	NA	NA	China: Hebei Province	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
11	AY843527.2	NA	Gallus gallus breed Silkies	China: Tianjin	55.8	51.7	53.7	24.53	32.08	31.37	29.47	0.782
12	AY999018.1	NA	NA	China: Shandong Province	55.3	50.8	53.4	24.53	31.13	32.35	29.47	0.781
13	DQ124934.1	NA	NA	China:Beijing	55.8	51.7	53.4	24.76	32.38	31.37	29.47	0.78
14	DQ124935.1	NA	NA	China:Anhui Province	55.5	51.7	54	24.53	32.08	31.37	29.47	0.775
15	DQ124936.1	NA	NA	China:Anhui Province	55.6	51.7	54	24.53	32.08	31.07	29.17	0.782
16	DQ141670.1	NA	NA	China: Shanghai	55.5	50.8	53.4	24.53	31.13	32.35	29.47	0.781
17	DQ141671.1	NA	NA	China: Shanghai	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.782
18	DQ141672.1	NA	NA	China: Henan Province	55	50.8	53.4	24.53	31.13	32.35	29.47	0.781
19	DQ141673.1	NA	NA	China: Shandong Pronvince	55.2	50.8	53.4	24.53	31.13	32.35	29.47	0.781
20	DQ991394.1	NA	NA	USA	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.782
21	EF176599.1	NA	NA	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
22	FJ172347.1	NA	broiler chicken	China	55.7	50.8	53.4	24.53	31.13	32.35	29.47	0.781
23	JF507715.1	NA	Gallus gallus	South Korea	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
24	JQ690762.1	Feces	Homo sapiens	China	55.2	51.7	54	24.53	32.08	31.37	29.47	0.776
25	JX260426.1	NA	Gallus gallus	China	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
26	JX964755.1	NA	Gallus gallus	China	55.4	52.5	54	24.53	32.08	30.39	30.53	0.777
27	KC414026.1	NA	Cat	China	55.8	52.5	53.7	24.53	32.08	30.39	30.53	0.786
28	KF224926.1	NA	Gallus gallus	China: Guangzhou	55.2	50.4	53.4	25.71	32.38	31.68	27.66	0.78
29	KF224927.1	NA	Gallus gallus	China: Guangzhou	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.782
30	KF224928.1	NA	Gallus gallus	China: Guangzhou	55	50.8	54	25	33.65	32	26.88	0.783
31	KF224929.1	NA	Gallus gallus	China: Guangzhou	55.1	51.7	53.7	24.53	32.08	31.37	29.47	0.782
32	KF224930.1	NA	Gallus gallus	China: Guangzhou	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
33	KF224931.1	NA	Gallus gallus	China: Guangzhou	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
34	KF224932.1	NA	Gallus gallus	China: Guangzhou	55.3	52.5	54	24.53	33.02	30.39	29.47	0.784
35	KF224933.1	NA	Gallus gallus	China: Guangzhou	54.9	51.7	53.7	25.23	31.78	30.69	29.79	0.784
36	KF224934.1	NA	Gallus gallus	China: Guangzhou	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
37	KF224935.1	NA	Gallus gallus	China: Guangzhou	55.3	51.7	53.7	24.53	32.08	31.68	29.79	0.781
38	KF224936.1	NA	Gallus gallus	China: Guangzhou	55.2	52.5	53.7	24.76	32.38	30.39	30.53	0.782
39	KF224937.1	NA	Gallus gallus	China: Guangzhou	55.4	51.7	53.7	24.53	32.08	31.37	29.79	0.782
40	KF224938.1	NA	Gallus gallus	China: Guangzhou	55.3	51.7	54	25.47	31.13	30.39	30.53	0.775
41	KJ872513.1	thymus	Gallus gallus; commercial broiler	Argentina	54.9	51.7	53.7	24.53	32.08	31.37	29.47	0.782
42	KJ872514.1	thymus	Gallus gallus; commercial broiler	Argentina	55.1	51.7	53.7	24.53	32.08	31.37	29.47	0.782
43	KU845734.1	NA	Gallus gallus	China	55.4	50.8	53.4	24.53	31.13	32.67	29.79	0.774
44	KU845735.1	NA	Gallus gallus	China	55.2	50.8	53.2	24.53	31.13	32.67	29.79	0.781
45	KU221054.1	NA	Gallus gallus	China	55.2	51.7	54	24.53	32.08	31.07	29.17	0.783
46	KU050677.1	NA	Gallus gallus	NA	55.1	51.7	53.7	24.53	32.08	31.37	29.47	0.782
47	KU050678.1	NA	Gallus gallus	NA	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
48	KU050679.1	NA	Gallus gallus	NA	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
49	KU050680.1	NA	Gallus gallus	NA	55.2	50.4	53.4	25.71	32.38	31.68	27.66	0.78
50	KU598851.1	NA	Gallus gallus	China	54.6	51.7	53.7	24.53	32.08	31.37	29.47	0.782
51	KU641013.1	NA	Gallus gallus	China	55.3	51.7	53.2	24.76	32.38	31.68	29.79	0.787
52	KU641014.1	NA	Gallus gallus	China	55.2	51.7	54	24.53	32.08	31.07	29.17	0.783
53	KU641015.1	NA	Gallus gallus	China	55.7	51.7	53.4	24.53	32.08	31.37	29.79	0.773
54	KU645506.1	NA	Gallus gallus	China	55.4	51.7	53.4	24.53	32.08	31.68	29.79	0.784
55	KU645507.1	NA	Gallus gallus	China	55	51.7	53.7	24.53	32.08	31.37	29.47	0.782
56	KU645508.1	NA	Gallus gallus	China	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.782
57	KU645509.1	NA	Gallus gallus	China	55.5	51.7	53.7	24.76	32.38	31.37	29.17	0.78
58	KU645510.1	NA	Gallus gallus	China	55.5	51.7	53.4	24.76	32.38	31.37	29.47	0.776
59	KU645511.1	NA	Gallus gallus	China	55.7	51.7	53.4	24.76	32.38	31.37	29.47	0.776
60	KU645512.1	NA	Gallus gallus	China	55.6	51.7	53.7	24.3	31.78	31.37	29.47	0.784
61	KU645513.1	NA	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
62	KU645514.1	NA	Gallus gallus	China	55.2	51.7	54	24.53	32.08	31.07	29.17	0.783
63	KU645515.1	NA	Gallus gallus	China	55.5	52.5	54	24.3	32.71	30.69	29.79	0.784
64	KU645516.1	NA	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.783
65	KU645517.1	NA	Gallus gallus	China	55.3	50.8	53.4	24.53	31.13	32.67	29.79	0.774
66	KU645518.1	NA	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782

(*continued on next page*)

Table 1 (*continued*)

(C)

Sr. No.	Accession Number	Tissue	Host	Origin	GC %	GC3s %	Nc	Mononucleotide Frequencies (%)				CAI
								T3 s	C3s	A3s	G3 s	
67	KU645519.1	NA	Gallus gallus	China	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
68	KU645520.1	NA	Gallus gallus	China	55.8	51.7	53.7	24.53	32.08	31.37	29.47	0.782
69	KU645521.1	NA	Gallus gallus	China	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
70	KU645522.1	NA	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
71	KU645523.1	NA	Gallus gallus	China	55.6	51.7	53.4	25	32.69	31.37	29.47	0.775
72	KU645524.1	NA	Gallus gallus	China	55.2	50.8	53.4	24.53	31.13	32.67	29.79	0.774
73	KU645525.1	NA	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.783
74	KY024579.1	feces	Gallus gallus	Brazil	54.9	51.7	53.7	24.53	32.08	31.37	29.47	0.782
75	KX811526.1	NA	Gallus gallus	China	55.4	50.8	53.4	24.53	31.13	32.67	29.79	0.774
76	KX447633.1	liver	Gallus gallus	China	55.2	51.7	53.4	24.53	32.08	31.68	29.79	0.782
77	KX447634.1	liver	Gallus gallus	China	55.5	51.7	53.4	24.53	32.08	31.68	29.79	0.782
78	KX447635.1	liver	Gallus gallus	China	55.2	51.7	53.4	24.53	32.08	32	29.79	0.782
79	KX447636.1	liver	Gallus gallus	China	55.3	51.7	53.7	24.76	32.38	31.37	29.17	0.78
80	KX447637.1	liver	Gallus gallus	China	55	51.7	54	24.53	32.08	31.37	29.17	0.779
81	KY053900.1	NA	Gallus gallus	India	55.9	51.7	53.7	24.53	32.08	31.37	29.47	0.782
82	KY486136.1	liver; spleen	Gallus gallus	China	55.5	51.7	53.7	24.53	32.08	31.37	29.47	0.782
83	KY486137.1	liver; spleen	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
84	KY486138.1	liver; spleen	Gallus gallus	China	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
85	KY486139.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
86	KY486140.1	liver; spleen	Gallus gallus	China	55.3	51.7	54	24.53	32.08	31.07	29.17	0.78
87	KY486141.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
88	KY486142.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
89	KY486143.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.779
90	KY486144.1	liver; spleen	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
91	KY486145.1	liver; spleen	Gallus gallus	China	55.3	51.7	53.7	24.53	32.08	31.37	29.47	0.782
92	KY486146.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
93	KY486147.1	liver; spleen	Gallus gallus	China	55.2	51.7	53.4	24.53	32.08	31.68	29.79	0.782
94	KY486148.1	liver; spleen	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
95	KY486149.1	liver; spleen	Gallus gallus	China	55.1	51.7	53.7	24.53	32.08	31.37	29.47	0.782
96	KY486150.1	liver; spleen	Gallus gallus	China	55.1	51.7	53.7	24.53	32.08	31.37	29.47	0.782
97	KY486151.1	liver; spleen	Gallus gallus	China	55.3	51.7	54	24.53	32.08	31.07	29.17	0.78
98	KY486152.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
99	KY486153.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.782
100	KY486154.1	liver; spleen	Gallus gallus	China	55.4	51.7	53.7	24.53	32.08	31.37	29.47	0.779
101	KY486155.1	liver; spleen	Gallus gallus	China	55.2	51.7	53.7	24.53	32.08	31.37	29.47	0.782
102	MG846491.1	feces	Gallus gallus	Brazil	54.8	51.7	53.7	24.53	32.08	31.37	29.47	0.782
103	AB027470.1	NA	wild type	NA	55.1	51.7	53.4	24.53	32.08	31.68	29.79	0.782
104	AJ297684.2	NA	NA	Germany:Cuxhaven	55.8	51.7	53.4	24.76	32.38	31.07	29.47	0.779
105	AJ297685.2	NA	NA	Germany:Cuxhaven	55.9	51.7	53.4	24.76	32.38	31.37	29.47	0.786
	MEAN				55.3	51.6	53.6	24.59	32.04	31.44	29.48	0.781

VP3) could be preferred (Table 2). 14 out of 18 preferred codons in VP1, 11 preferred codons in VP2 and 10 preferred codons in VP3 were C/G-ended. For VP1, out of the 6 preferred codons which were over-represented (RSCU > 1.6), 5 were C/G-ended while 1 was A/U-ended. For VP2, 5 out of 10 over-represented preferred codons were C/G-ended while for VP3, 6 out of 12 preferred over-represented codons were C/G-ended.

3.4. Codon usage bias

ENC values were calculated to quantify the extent of codon usage bias within coding sequences of the viral protein genes of CAV (Table 1). ENC values for VP1 ranged from 52.58 to 57.42 with a mean of 55 and SD of 0.93. ENC values for VP2 ranged from 50.26 to 53.23 with a mean of 51.76 and SD of 0.53. ENC values for VP3 ranged from

Table 2
Predominantly used codons in CAV according to RSCU analysis and comparison with codon usage of its host, *Gallus gallus*. The codon usage pattern of CAV was a mix of coincidence and antagonism with respect to its host.

AA	Predominantly used codon			
	VP1	VP2	VP3	*Gallus gallus*
Phe	UUU	UUC	UUC	UUC
Leu	CUC	CUU	CUA	CUG
Ile	AUC	AUC	AUC	AUC
Val	GUG	GUA	GUG	GUG
Ser	AGC	UCA	AGC	AGC
Pro	CCC	CCG	CCC	CCA
Thr	ACA	ACC	ACU	ACA
Ala	GCG	GCC	GCU	GCU
Tyr	UAC	UAC	UAC	UAC
His	CAC	CAC	CAC	CAC
Gln	CAA	CAA	CAA	CAG
Asn	AAC	AAC	AAU	AAC
Lys	AAG	AAG	AAG	AAG
Asp	GAC	GAC	GAC	GAU
Glu	GAG	GAA	GAA	GAG
Cys	UGC	UGC	UGC	UGC
Arg	AGA	AGA	AGA	AGA
Gly	GGC	GGA	GGU	GGC

53.2 to 54 with a mean of 53.66 and SD of 0.19. ENC values for all the three genes were below the expected values, suggesting the prevalence of other factors such as natural selection in addition to mutational pressure. With mean ENC values lying between 51 and 55, all three genes showed slight codon bias; however, the lowest ENC values of VP2 among the three, suggested slightly greater codon bias than the two other genes.

3.5. Codon usage adaptation

The Codon Adaptation Index (CAI) is often used as a measure for codon usage optimization and adaptation of viral genes to its host. CAI values for VP1 ranged from 0.75 to 0.783 with a mean value of 0.766 and SD of 0.007. CAI values for VP2 ranged from 0.772 to 0.785 with a mean value of 0.777 and SD of 0.002, while that for VP3 ranged from 0.773 to 0.787 with a mean value of 0.781 and SD of 0.002. High values of CAI represent a higher proportion of the most abundant codons and expression level in the host. High CAI values using codon usage of *Gallus gallus* as a reference dataset suggests the adaptation of the virus to the host.

3.6. The relation between codon usage pattern of CAV and its host

The codon usage of CAV and *Gallus gallus* was compared in order to analyse the effect of the host on its pattern (Table 2). The highest number of preferred viral codons coincident with *Gallus gallus*, was observed in VP1 (12 out of 18 preferred codons). For VP2 and VP3, 8 and 10 out of 18 preferred codons, respectively, were coincident with that of *Gallus gallus*. The viral codon usage pattern was not entirely coincident with that of the host *Gallus gallus*. These results combined with CAI values, possibly indicate the evolution of the virus so as to adapt to its host for its survival and proliferation.

3.7. Trends in codon usage variation

Codon usage data was analysed using multivariate statistical techniques such as correspondence analysis (COA). To determine the trends in the codon usage variation of different CAV isolates, we performed a COA on the RSCU values which were examined as a single data set

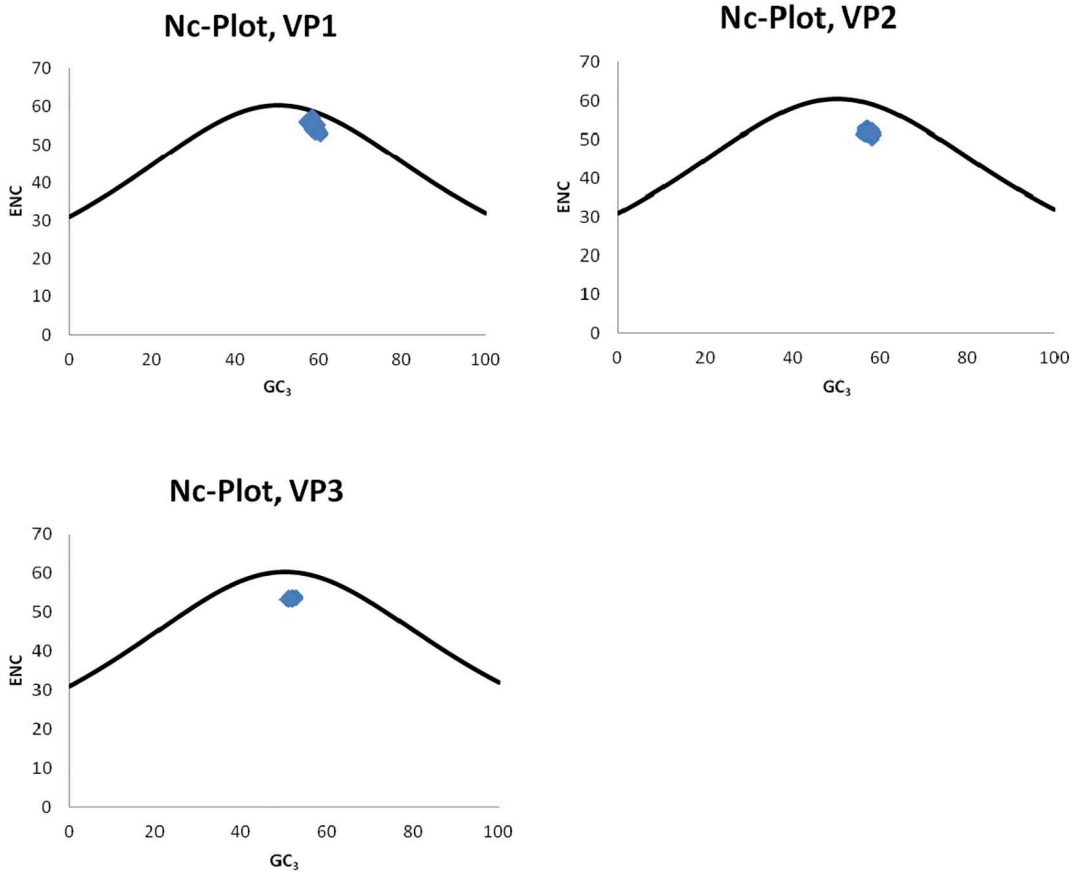

Fig. 2. The relationship between the ENC and GC_3 of VP1, VP2 and VP3 genes of chicken anaemia virus. The sequences clustered closely below the expected curve, for all three genes, with ENC values > 50 for all of them.

based on the RSCU values of VP1, VP2 and VP3. For VP1, the first principal axis (f'1) accounted for 36.77% of the total variation, while second principal axes (f'2) accounted for 18.09% of the total variation respectively. For VP2, f'1 accounted for 39.13% of the total variation, while f'2, accounted for 32.49% of the total variation respectively. For VP3, the first two axes responsible for maximum variation accounted for 28.97% and 22.48% of the total variation respectively. Further correlation analysis of these axes with different genomic properties was performed. The trends correlating with the principal axis responsible for maximum variation helped in elucidating the extent to which mutational pressure and natural selection contribute, as mentioned in the following results.

3.8. Correlation analysis between GC_3 and ENC values

Analysis of Nc plot, which is the curve plotted between ENC and GC_3, aids in studying codon usage variation. All the points aggregated below the expected ENC curve, slightly on the right, indicating mutational pressure to play a dominant role in shaping the codon usage patterns of CAV genes, in addition to other factors (Fig. 2). Furthermore, Pearson correlation analysis also showed that GC_3 values for all three genes were correlated with ENC values ($p < 0.05$).

3.9. Role of mutation pressure in the shaping of CAV codon usage patterns

In addition to the Nc plot, correlation analysis between different nucleotide constraints, ENC, GC, GC_3 and the CoA axes were performed to assess the extent to which mutation pressure affects the codon usage patterns of CAV genes. Several strong and significant correlations were observed ($p < 0.05$). The GC and GC_3 were positively correlated with high significance, for all three genes VP1, VP2 and VP3 (Table 3). It was observed that GC and ENC values, as well as GC_3 and ENC values, were significantly positively and negatively correlated respectively, for all three genes (Table 3); while AU_3 and ENC values were significantly positively correlated for all three genes (data not shown). Furthermore, both GC3 and ENC values were correlated with the first principal axis (f'1) while only ENC values were correlated with f'1 for VP1 and VP3. $\%T_3$ and $\%C_3$ were correlated with the first principal axis (f'1) for VP2, but not for VP1 and VP3 while $\%A_3$ and $\%G_3$ were correlated with the second principal axis (f'2) for VP1 and VP3, but not for VP2. These correlations of nucleotide constraints, GC_3 and ENC values with principal axes and that between GC_3 and ENC indicate the significant role of mutational pressure in shaping the codon usage of CAV. The results also indicate that mutational pressure contributes to a greater extent to the codon usage in VP2 as compared to VP1 and VP3, owing to high nucleotide constraints.

Table 4
Dinucleotide odds ratio in CAV sequences.

	VP1	VP2	VP3
TT	1.315	1.471	0.916
TC	1.005	1.071	1.196
TA	0.867	0.669	0.565
TG	1.026	0.982	1.238
CT	1.081	1.133	1.286
CC	0.943	0.992	1.078
CA	1.123	0.923	0.860
CG	0.890	0.964	0.956
AT	1.032	0.928	0.896
AC	1.027	0.936	0.884
AA	0.924	1.103	1.143
AG	1.054	1.091	1.044
GT	0.789	0.662	0.734
GC	1.021	0.990	0.999
GA	1.083	1.266	1.331
GG	0.940	0.932	0.766

3.9.1. Influence of relative dinucleotide abundance on the codon usage bias

The earlier study suggested that relative dinucleotide abundance can affect the overall codon usage bias (Karlin and Burge, 1995). Relative abundance of 16 dinucleotides was calculated to study the possible effect of dinucleotides on codon usage in CAV genes (Table 4). For VP1, only TpT showed a deviation from normal range and was over-represented. However, none of the codons containing UpU were over-represented or preferred codons, except for UUU as a preferred codon. For VP2, TpT was over-represented while TpA and GpT were under-represented. However, none of the codons containing UpU were over-represented or preferred codons, except for CUU as a preferred codon. Two out of six UpA codons were under-represented (UUA and CUA) and two out of six UpA codons (UAC and GUA) were preferred; four out of eight GpU codons were under-represented and only GpU (GUA) was preferred codon for VP2. For VP3, TpG, CpT and GpA were over-represented while TpA, GpT and GpG were under-represented. Eight out of eighteen preferred codons for VP3 consisted of UpG, CpU or GpA while three preferred codons consisted of UpA, GpU or GpG.

3.9.2. Effect of natural selection in shaping the codon usage patterns of CAV

The GRAVY and AROMO values were calculated and Pearson correlation analysis was performed between GRAVY and AROMA values (Table 5). For VP1, GRAVY was significantly correlated with f'2 while AROMO was significantly correlated with f'1. GRAVY showed significant correlation with GC, GC_3 and ENC. For VP2, GRAVY was not correlated with any of the axes, but AROMO was negatively correlated with f'2. GRAVY was significantly correlated to GC, but neither to GC_3

Table 3
Results of correlation analysis. The values in table denote correlation coefficient, 'r' ($p < 0.05$). NS = Non-Significant. f'1 and f'2 denote first two principal axes of Correspondence Analysis (CoA) respectively. The results demonstrate the effect of mutation pressure in shaping the codon usage pattern of CAV.

	%T3	%C3	%A3	%G3	ENC	%GC3	%GC
VP1							
f'1	NS	NS	NS	NS	0.5029	NS	NS
f'2	−0.26	NS	0.4496	−0.3684	0.5448	NS	NS
%GC	−0.5579	0.5498	−0.6555	0.5927	−0.6079	0.92	–
ENC	0.3155	−0.3092	0.4589	−0.4291	–	−0.5943	−0.6079
VP2							
f'1	0.4795	−0.2957	NS	NS	−0.3402	−0.2100	NS
f'2	NS	NS	NS	NS	NS	NS	NS
%GC	−0.4630	0.5113	−0.7562	0.41	−0.5061	0.8251	–
ENC	0.1989	−0.3764	0.4449	−0.22	–	−0.4817	−0.5061
VP3							
f'1	NS	NS	NS	NS	−0.3202	NS	NS
f'2	NS	NS	−0.2665	0.29	0.4500	0.28	NS
%GC	NS	0.3348	−0.6522	NS	NS	0.49	–
ENC	0.2014	NS	NS	NS	–	−0.1892	NS

Table 5
Results of correlation analysis. The values in table denote correlation coefficient, 'r' ($p < 0.05$). NS = Non-Significant. f'1 and f'2 denote first two principal axes of Correspondence Analysis (CoA) respectively. The results demonstrate the effect of natural selection in shaping the codon usage pattern of CAV.

	f'1	f'2	%GC	%GC3	Nc
VP1					
GRAVY	NS	−0.3479	0.2993	0.2303	−0.3475
AROMO	−0.2790	NS	NS	NS	−0.2224
VP2					
GRAVY	NS	NS	0.2792	NS	NS
AROMO	NS	−0.5676	NS	NS	NS
VP3					
GRAVY	NS	NS	0.3039	0.2193	−0.4192
AROMO	NS	NS	NS	NS	NS

nor to ENC. For VP3, neither AROMO nor GRAVY was correlated to any of the axes. GRAVY was significantly correlated to GC, GC_3 and ENC. The results showed that natural selection had played only a minor role in the codon usage patterns of CAV genes owing to the absence of strong correlations with the first principal axis. The results also indicate that natural selection played the least role in codon usage of VP2 as compared to the other two genes; whereas mutational pressure was dominant in VP2 as compared to other two genes as observed in the previous section.

4. Discussion

Nucleotide constraints have an impact on the codon usage, especially at the third position of the codon and to elucidate these effects nucleotide composition analysis in CAV was done. The nucleotide composition analysis suggested the presence of compositional constraint in the CAV genes. RSCU analysis performed herein aided us in assessing the patterns of synonymous codon usage in CAV proteins. Combining these analyses assists us in deciphering whether the selection of preferred codons has been influenced by mutational pressure or natural selection.

A mean ENC value of > 40 suggests a relatively stable and conserved sequence among different CAV isolates. Gene expression and ENC values are inversely related (Wright, 1990). Our analysis shows that with ENC values > 50, the overall codon usage bias and gene expression among different CAV isolates is low and slightly biased, for all three viral protein genes. Similar results were obtained upon codon bias analysis of small DNA viruses such as torque teno virus (TTSuV) (Zhang et al., 2013b; Zhang et al., 2013c) and porcine circovirus (PCV) (Chen et al., 2014). It has been reported that CAI value defines the codon usage bias and efficiency of protein expression. Our results show a high mean CAI value for all the CAV genes, indicating high codon usage bias and higher efficiency of protein expression in the host.

It has been postulated that an Nc-plot of genes, whose codon choice is constrained only by a $G_3 + C_3$ mutational bias, will lie on or just below the continuous curve of the predicted ENC values (Wright, 1990). Our result showed that all the points aggregated below the expected ENC curve, slightly on the right, thus indicating that the codon usage in the CAV genes is mostly influenced by mutational pressure. Mutation pressure and natural selection are considered the two major factors that shape codon usage patterns (Tatarinova et al., 2010). Our results indicate that compositional constraints under mutation pressure have shaped the codon usage pattern for CAV genes to a major extent.

It can be expected that the codon usage pattern of viruses would be affected by its hosts to some extent due to its obligate parasitism (Zhou et al., 2005). For instance, codon usage pattern of poliovirus is reported to be mostly coincident with its host (Mueller et al., 2006), while the codon usage pattern of Hepatitis A virus is reported to be antagonistic with its host (Sanchez et al., 2003). It was seen that while the codon usage in CAV was coincident with that of *Gallus gallus* in some parts, it was asynchronous in others. These results imply that natural selection also influences CAV codon usage patterns. These results are further in accordance with comparatively high CAI values, thus reflecting that this interplay of codon usage has influenced viral fitness and survival of CAV in its hosts and consequent efficiency of protein synthesis. Similarly, the codon usage pattern of TTSuv1 was observed to be more similar to its host as compared to TTSuv2 and was linked to its evasion of antiviral cell response and its adaptation to the host (Zhang et al., 2013b).

Performing correspondence analysis on the RSCU values revealed the major trends influencing the codon usage variation. The correlation of these axes with different analysis values enabled us a step further in elucidating the role of mutational pressure and natural selection on codon usage.

It has been suggested that the frequencies of nucleotides A and U/T should be equal to that of C and G at the third position of the codon if mutational pressure affects the synonymous codon usage bias (Zhang et al., 2013d). However, variations in the nucleotide base composition were noted in case of CAV (Table 1), indicating that synonymous codon usage bias could also be influenced by other factors such as natural selection. The correlation of GRAVY and AROMO with the principal axes indicated that although natural selection has influenced codon usage of CAV protein genes to some extent, it is much lesser compared to mutational pressure.

Codon usage bias can potentially influence the level of protein expression, and synthetically optimizing codon might, therefore, be beneficial for cross-species expression. The expressed proteins could be obtained at higher levels by employing codon optimization in a prokaryotic and eukaryotic expression system. In spite of its many other benefits, DNA vaccine is limited in its application by its insufficient immunogenicity. One promising approach for enhancing its immunogenicity is to maximize its expression in the immunized host. It is well-established that the optimization of the codons in the construction of DNA vaccines could increase its efficiency (Chen et al., 2008; Ingolotti et al., 2010; Yan et al., 2007; Zhu et al., 2010). Another approach could be to generate an attenuated virus particle with deoptimized codon usage, thereby downregulating its protein synthesis (Mueller et al., 2006). Changes in the number of CpG and UpA dinucleotides have been recognized as yet another major driving force in codon bias and thereby modulating immunity and protein expression, and has been discussed extensively (Martínez et al., 2016; Vabret et al., 2017). Introduction of CpG and omittance of UpA also results in optimization of various mammalian genes and their elevated expression (Fath et al., 2011). It was also explored for genetic inactivation of poliovirus infectivity by increasing CpG and UpA frequencies based on the immune response elicited by CpG (Burns et al., 2009). Moreover, the implementation of above approaches has a great impact in designing therapeutics against human diseases as well (Mauro and Chappell, 2014).

Although the phylogenetic tree showed three distinct clusters, no conclusion could be drawn regarding the correspondence between clustering and the geographical origins of the samples. A major reason behind this could be that most isolates are from the same province and hence significant correlation cannot be obtained.

The evolutionary adaptation of viruses to their hosts also includes the codon usage in its genome in agreement with the codon preferences of the host. Virus-host interactions may be regulated by altering the codon usage of viral genes based on the appropriate information from the host genome. Studies analysing the codon composition and the potential effects of its alterations are being conducted in polioviruses (Coleman et al., 2008; Mueller et al., 2006), parvoviruses (Zhi et al., 2010), influenza virus (Tenbusch et al., 2010), HIV (Ramakrishna et al., 2004), respiratory syncytial virus (Kim et al., 2010b), papillomaviruses (Kim and Sin, 2005), among others.

To the best of our knowledge, this is the first report of codon usage analysis in CAV genes and is expected to deepen our understanding of its codon usage patterns. The analysis suggested that all the genes demonstrate a low codon usage bias, however, the bias is slightly greater in the VP2 gene as compared toVP1 and VP3. The major factor contributing to the codon bias in CAV is mutational pressure followed by the influence of hosts. In addition, contributions by other factors such as natural selection has also influenced the codon usage. Further, studies can be pursued based on codon optimization of the CAV protein genes for its maximum expression. Our study could, therefore, aid in designing efficient vaccines.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The CAV research in our laboratory is currently supported by the internal grant from the Indian Institute of Technology Guwahati.

References

AboElkhair, M., Abd El-Razak, A.G., Metwally, A.E., 2014. Molecular characterization of chicken anemia virus circulating in chicken flocks in Egypt. Adv. Virol. 797151.
Adair, B.M., 2000. Immunopathogenesis of chicken anemia virus infection. Develop. Comp. Immunol. 24, 247–255.
Bhatt, P., Shukla, S.K., Mahendran, M., Dhama, K., Chawak, M.M., Kataria, J.M., 2011. Prevalence of chicken infectious anaemia virus (CIAV) in commercial poultry flocks of northern India: a serological survey. Transbound. Emerg. Dis. 58, 458–460.
Burns, C.C., Campagnoli, R., Shaw, J., Vincent, A., Jorba, J., Kew, O., 2009. Genetic inactivation of poliovirus infectivity by increasing the frequencies of CpG and UpA dinucleotides within and across synonymous capsid region codons. J. Virol. 83, 9957–9969.
Cao, H.W., Li, D.S., Zhang, H., 2014. Analysis of synonymous codon usage in Newcastle disease virus hemagglutinin-neuraminidase (HN) gene and fusion protein (F) gene. Virus Disease 25, 132–136.
Chen, M.W., Cheng, T.J., Huang, Y., Jan, J.T., Ma, S.H., Yu, A.L., Wong, C.H., Ho, D.D., 2008. A consensus-hemagglutinin-based DNA vaccine that protects mice against divergent H5N1 influenza viruses. Proc. Natl. Acad. Sci. U. S. A. 105, 13538–13543.
Chen, Y., Sun, J., Tong, X., Xu, J., Deng, H., Jiang, Z., Jiang, C., Duan, J., Li, J., Zhou, P., Wang, C., 2014. First analysis of synonymous codon usage in porcine circovirus. Arch. Virol. 159, 2145–2151.
Coghlan, A., Wolfe, K.H., 2000. Relationship of codon bias to mRNA concentration and protein length in Saccharomyces cerevisiae. Yeast (Chichester, England) 16, 1131–1145.
Coleman, J.R., Papamichail, D., Skiena, S., Futcher, B., Wimmer, E., Mueller, S., 2008. Virus attenuation by genome-scale changes in codon pair bias. Science 320, 1784–1787.
Comeron, J.M., Aguade, M., 1998. An evaluation of measures of synonymous codon usage bias. J. Mol. Evol. 47, 268–274.
Craig, M.I., Rimondi, A., Delamer, M., Sansalone, P., Konig, G., Vagnozzi, A., Pereda, A., 2009. Molecular characterization of chicken infectious anemia virus circulating in Argentina during 2007. Avian Dis. 53, 331–335.
Duret, L., Mouchiroud, D., 1999. Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. Proc. Natl. Acad. Sci. 96, 4482–4487.
Engstrom, B.E., 1999. Prevalence of antibody to chicken anaemia virus (CAV) in Swedish chicken breeding flocks correlated to outbreaks of blue wing disease (BWD) in their progeny. Acta Vet. Scand. 40, 97–107.
Epstein, R.J., Lin, K., Tan, T.W., 2000. A functional significance for codon third bases. Gene 245, 291–298.
Eregae, M.E., Dewey, C.E., McEwen, S.A., Ouckama, R., Ojkic, D., Guerin, M.T., 2014. Flock prevalence of exposure to avian adeno-associated virus, chicken anemia virus, fowl adenovirus, and infectious bursal disease virus among Ontario broiler chicken flocks. Avian Dis. 58, 71–77.
Fath, S., Bauer, A.P., Liss, M., Spriestersbach, A., Maertens, B., Hahn, P., Ludwig, C., Schäfer, F., Graf, M., Wagner, R., 2011. Multiparameter RNA and codon optimization: a standardized tool to assess and enhance autologous mammalian gene expression. PLoS One 6 (3), e17596.
Firth, G., Imai, K., 1990. Isolation of chicken anaemia agent from Australian poultry. Aust. Vet. J. 67, 301–302.
Gholami-Ahangaran, M., 2015. Serological and molecular detection of chicken anaemia virus in Iranian poultry flocks. Vet. Ital. 51, 211–215.
Goni, N., Iriarte, A., Comas, V., Sonora, M., Moreno, P., Moratorio, G., Musto, H., Cristina, J., 2012. Pandemic influenza a virus codon usage revisited: biases, adaptation and implications for vaccine strain development. Virol. J. 9, 263.
Grantham, R., Gautier, C., Gouy, M., 1980a. Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type. Nucleic Acids Res. 8, 1893–1912.
Grantham, R., Gautier, C., Gouy, M., Mercier, R., Pave, A., 1980b. Codon catalog usage and the genome hypothesis. Nucleic Acids Res. 8, r49–r62.
Gu, W., Zhou, T., Ma, J., Sun, X., Lu, Z., 2004. The relationship between synonymous codon usage and protein structure in Escherichia coli and Homo sapiens. Biosystems 73, 89–97.
Hassan, S., Mahalingam, V., Kumar, V., 2009. Synonymous codon usage analysis of thirty two mycobacteriophage genomes. Adv. Bioinforma. 316936.
Ingolotti, M., Kawalekar, O., Shedlock, D.J., Muthumani, K., Weiner, D.B., 2010. DNA vaccines for targeting bacterial infections. Expert Rev. Vaccines 9, 747–763.
Jenkins, G.M., Holmes, E.C., 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 92, 1–7.
Karlin, S., Burge, C., 1995. Dinucleotide relative abundance extremes: a genomic signature. Trends Genet. 11, 283–290.
Kim, M.S., Sin, J.I., 2005. Both antigen optimization and lysosomal targeting are required for enhanced anti-tumour protective immunity in a human papillomavirus E7-expressing animal tumour model. Immunology 116, 255–266.
Kim, H.R., Kwon, Y.K., Bae, Y.C., Oem, J.K., Lee, O.S., 2010a. Molecular characterization of chicken infectious anemia viruses detected from breeder and broiler chickens in South Korea. Poult. Sci. 89, 2426–2431.
Kim, S., Jang, J.E., Yu, J.R., Chang, J., 2010b. Single mucosal immunization of recombinant adenovirus-based vaccine expressing F1 protein fragment induces protective mucosal immunity against respiratory syncytial virus infection. Vaccine 28, 3801–3808.
Koch, G., van Roozelaar, D.J., Verschueren, C.A., van der Eb, A.J., Noteborn, M.H., 1995. Immunogenic and protective properties of chicken anaemia virus proteins expressed by baculovirus. Vaccine 13, 763–770.
Kumar, C.S., Kumar, S., 2014. Species based synonymous codon usage in fusion protein gene of Newcastle disease virus. PLoS One 9, e114754.
Kumar, C.S., Kumar, S., 2017. Synonymous codon usage of genes in polymerase complex of Newcastle disease virus. J. Basic Microbiol. 57, 481–503.
Kumar, C.S., Hazarika, N.M., Kumar, S., 2015. Analysis of synonymous codon usage in the VP2 protein gene of infectious bursal disease virus. Arch. Virol. 160, 2359–2366.
Kye, S.J., Kim, J.Y., Seul, H.J., Kim, S., Kim, S.E., Lee, H.S., Sorn, S., Choi, K.S., 2013. Phylogenetic analysis and genetic characterization of chicken anemia virus isolates from Cambodia. Poult. Sci. 92, 2681–2686.
Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.
Lai, G.-H., Lin, M.-K., Lien, Y.-Y., Cheng, J.-H., Sun, F.-C., Lee, M.-S., Chen, H.-J., Lee, M.-S., 2018. Characterization of the DNA binding activity of structural protein VP1 from chicken anaemia virus. BMC Vet. Res. 14, 155.
Li, Z.P., Ying, D.Q., Li, P., Li, F., Bo, X.C., Wang, S.Q., 2010. Analysis of synonymous codon usage bias in 09H1N1. Virol. Sin. 25, 329–340.
Los, M., Panigrahi, S., Rashedi, I., Mandal, S., Stetefeld, J., Essmann, F., Schulze-Osthoff, K., 2009. Apoptin, a tumor-selective killer. Biochim. et Biophys. Acta (BBA) 1793, 1335–1342.
Makhija, A., Kumar, S., 2015. Analysis of synonymous codon usage in spike protein gene of infectious bronchitis virus. Can. J. Microbiol. 61, 983–989.
Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.
Martínez, M.A., Jordan-Paiz, A., Franco, S., Nevot, M., 2016. Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. Trends Microbiol. 24, 134–147.
Mauro, V.P., Chappell, S.A., 2014. A critical analysis of codon optimization in human therapeutics. Trends Mol. Med. 20, 604–613.
McConnell, C.D., Adair, B.M., McNulty, M.S., 1993. Effects of chicken anemia virus on cell-mediated immune function in chickens exposed to the virus by a natural route. Avian Dis. 37, 366–374.
McIlroy, S.G., McNulty, M.S., Bruce, D.W., Smyth, J.A., Goodall, E.A., Alcorn, M.J., 1992. Economic effects of clinical chicken anemia agent infection on profitable broiler production. Avian Dis. 36, 566–574.
McNulty, M.S., Connor, T.J., McNeilly, F., McLoughlin, M.F., Kirkpatrick, K.S., 1990. Preliminary characterisation of isolates of chicken anaemia agent from the United Kingdom. Avian Pathol. 19, 67–73.
McNulty, M.S., McIlroy, S.G., Bruce, D.W., Todd, D., 1991. Economic effects of subclinical chicken anemia agent infection in broiler chickens. Avian Dis. 35, 263–268.
Mueller, S., Papamichail, D., Coleman, J.R., Skiena, S., Wimmer, E., 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J. Virol. 80, 9687–9696.
Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28, 292.
Noteborn, M.H.M., 2004. Chicken anemia virus induced apoptosis: underlying molecular mechanisms. Vet. Microbiol. 98, 89–94.
Noteborn, M.H., de Boer, G.F., van Roozelaar, D.J., Karreman, C., Kranenburg, O., Vos, J.G., Jeurissen, S.H., Hoeben, R.C., Zantema, A., Koch, G., et al., 1991. Characterization of cloned chicken anemia virus DNA that contains all elements for the infectious replication cycle. J. Virol. 65, 3131–3139.
Noteborn, M.H., Todd, D., Verschueren, C.A., de Gauw, H.W., Curran, W.L., Veldkamp, S., Douglas, A.J., McNulty, M.S., Koch, G., 1994. A single chicken anemia virus protein induces apoptosis. J. Virol. 68, 346–351.
Olszewska-Tomczyk, M., Swieton, E., Minta, Z., Smietanka, K., 2016. Occurrence and Phylogenetic Studies of Chicken Anemia Virus from polish broiler Flocks. Avian Dis. 60, 70–74.
Peden, J., 1999. Analysis of Codon Usage. University of Nottingham, UK.
Peters, M.A., Jackson, D.C., Crabb, B.S., Browning, G.F., 2002. Chicken anemia virus VP2

is a novel dual specificity protein phosphatase. J. Biol. Chem. 277, 39566–39573.
Peters, M.A., Crabb, B.S., Washington, E.A., Browning, G.F., 2006. Site-directed mutagenesis of the VP2 gene of Chicken anemia virus affects virus replication, cytopathology and host-cell MHC class I expression. J Gen Virol 87, 823–831.
Puigbo, P., Bravo, I.G., Garcia-Vallve, S., 2008. CAIcal: a combined set of tools to assess codon usage adaptation. Biol. Direct 3, 38.
Ramakrishna, L., Anand, K.K., Mohankumar, K.M., Ranga, U., 2004. Codon optimization of the tat antigen of human immunodeficiency virus type 1 generates strong immune responses in mice following genetic immunization. J. Virol. 78, 9174–9189.
Rosario, K., Breitbart, M., Harrach, B., Segales, J., Delwart, E., Biagini, P., Varsani, A., 2017. Revisiting the taxonomy of the family Circoviridae: establishment of the genus Cyclovirus and removal of the genus Gyrovirus. Arch. Virol. 162, 1447–1463.
Rosenberger, J.K., Cloud, S.S., 1998. Chicken anemia virus. Poult. Sci. 77, 1190–1192.
Sanchez, G., Bosch, A., Pinto, R.M., 2003. Genome variability and capsid structural constraints of hepatitis a virus. J. Virol. 77, 452–459.
Schat, K.A., 2009. Chicken anemia virus. In: de Villiers, E.M., Hausen, H.Z. (Eds.), TT Viruses: The Still Elusive Human Pathogens. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 151–183.
Shackelton, L.A., Parrish, C.R., Holmes, E.C., 2006. Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. J. Mol. Evol. 62, 551–563.
Sharp, P.M., Li, W.H., 1986. An evolutionary perspective on synonymous codon usage in unicellular organisms. J. Mol. Evol. 24, 28–38.
Sharp, P.M., Li, W.H., 1987. The codon Adaptation Index–a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295.
Smuts, H.E., 2014. Novel Gyroviruses, including Chicken Anaemia Virus, in Clinical and Chicken Samples from South Africa. Adv Virol 2014, 321284.
Snoeck, C.J., Komoyo, G.F., Mbee, B.P., Nakoune, E., Le Faou, A., Okwen, M.P., Muller, C.P., 2012. Epidemiology of chicken anemia virus in Central African Republic and Cameroon. Virol. J. 9, 189.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tatarinova, T.V., Alexandrov, N.N., Bouck, J.B., Feldmann, K.A., 2010. GC3 biology in corn, rice, sorghum and other grasses. BMC Genomics 11, 308.
Tenbusch, M., Grunwald, T., Niezold, T., Storcksdieck Genannt Bonsmann, M., Hannaman, D., Norley, S., Uberla, K., 2010. Codon-optimization of the hemagglutinin gene from the novel swine origin H1N1 influenza virus has differential effects on CD4(+) T-cell responses and immune effector mechanisms following DNA electroporation in mice. Vaccine 28, 3273–3277.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Vabret, N., Bhardwaj, N., Greenbaum, B.D., 2017. Sequence-specific Sensing of Nucleic Acids. Trends Immunol. 38, 53–65.
Wong, E.H., Smith, D.K., Rabadan, R., Peiris, M., Poon, L.L., 2010. Codon usage bias and the evolution of influenza A viruses. Codon Usage Biases of Influenza Virus. BMC Evol. Biol. 10, 253.
Wright, F., 1990. The 'effective number of codons' used in a gene. Gene 87, 23–29.
Yan, J., Yoon, H., Kumar, S., Ramanathan, M.P., Corbitt, N., Kutzler, M., Dai, A., Boyer, J.D., Weiner, D.B., 2007. Enhanced cellular immune responses elicited by an engineered HIV-1 subtype B consensus-based envelope DNA vaccine. Mol. Ther. 15, 411–421.
Zhang, X., Liu, Y., Wu, B., Sun, B., Chen, F., Ji, J., Ma, J., Xie, Q., 2013a. Phylogenetic and molecular characterization of chicken anemia virus in southern China from 2011 to 2012. Sci. Rep. 3, 3519.
Zhang, Z., Dai, W., Dai, D., 2013b. Synonymous codon usage in TTSuV2: analysis and comparison with TTSuV1. PLoS One 8, e81469.
Zhang, Z., Dai, W., Wang, Y., Lu, C., Fan, H., 2013c. Analysis of synonymous codon usage patterns in torque Teno sus virus 1 (TTSuV1). Arch. Virol. 158, 145–154.
Zhang, Z., Dai, W., Wang, Y., Lu, C., Fan, H., 2013d. Analysis of synonymous codon usage patterns in torque Teno sus virus 1 (TTSuV1). Arch. Virol. 158, 145–154.
Zhi, N., Wan, Z., Liu, X., Wong, S., Kim, D.J., Young, N.S., Kajigaya, S., 2010. Codon optimization of human parvovirus B19 capsid genes greatly increases their expression in nonpermissive cells. J. Virol. 84, 13059–13062.
Zhou, H., Wang, H., Huang, L.F., Naylor, M., Clifford, P., 2005. Heterogeneity in codon usages of sobemovirus genes. Arch. Virol. 150, 1591–1605.
Zhu, Y., Lu, F., Dai, Y., Wang, X., Tang, J., Zhao, S., Zhang, C., Zhang, H., Lu, S., Wang, S., 2010. Synergistic enhancement of immunogenicity and protection in mice against Schistosoma japonicum with codon optimization and electroporation delivery of SjTPI DNA vaccines. Vaccine 28, 5347–5355.