Virus Research 353 (2025) 199528

Contents lists available at ScienceDirect

Virus Research

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

Analysis of synonymous codon usage bias of Lassa virus

Siddiq Ur Rahman [a,b,1], Yikui Hu [c,1], Hassan Ur Rehman [b,1], May M. Alrashed [d], Kotb A. Attia [e,*], Ubaid Ullah [b], Huiying Liang [a,*]

[a] Medical Big Data Center, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong Province 510080, PR China
[b] Department of Computer Science & Bioinformatics, Khushal Khan Khattak University, Karak, Khyber Pakhtunkhwa, 27200, Pakistan
[c] Department of Neurology, Wuhan Wuchang Hospital, Wuhan, PR China
[d] Department of Clinical Laboratory Science, College of Applied Medical Sciences, King Saud University, P.O. Box 2455, 11451 Riyadh, Saudi Arabia
[e] Department of Biochemistry, College of Science, King Saud University, P.O. Box 2455, 11451 Riyadh, Saudi Arabia

ARTICLE INFO

Keywords:
Lassa virus
Codon usage
Mutation pressure
Natural selection
Effective number of codons (ENC)

ABSTRACT

Lassa virus genome consists of two single-stranded, negative-sense RNA segments that lie in the genus *Arenavirus*. The disease associated with the Lassa virus is distributed all over the world, with approximately 3,000,000–5,000,000 infections diagnosed annually in West Africa. It shows high health risks to the human being. Previous research used the evolutionary time scale and adaptive evolution to describe the Lassa virus population pattern. However, it is still unclear how the Lassa virus takes advantage of synonymous codons. In this study, we analyzed the codon usage bias in 162 Lassa virus strains by calculating and comparing the nucleotide contents, effective number of codons (ENC), codon adaptation index (CAI), relative synonymous codon usage (RSCU), and others. The results disclosed that LASV strains are rich in A/T. The average ENC value indicated a low codon usage bias in LASVs. The ENC-plot, neutrality plot and parity rule 2 plot demonstrated that, besides mutational pressure, other factors like natural selection also contributed to codon usage bias. This study is significant because it described the pattern of codon usage in the genomes of the Lassa viruses and provided the information needed for a fundamental evolutionary study of them.

1. Introduction

The Lassa virus (LASV), which belongs to the family *Arenaviridae* and is the etiologic agent of viral hemorrhagic fever disease, infects between 3,00,000 and 5,00,000 people each year in West Africa (McCormick et al., 1987; McCormick and Fisher-Hoch, 2002). The segmented genome of the LASV is a two-single-stranded, negative-sense RNA virus. The LASV has a segmented genome specialized for distinct purpose: S (small), M (medium), and L (large) segments encode nucleocapsid (N), glycoproteins (Gn and Gc), and RNA-dependent polymerases (RdRp), respectively (Buchmeier et al., 1987; Lukashevich et al., 1984; Vieth et al., 2004).

Humans are mostly exposed to LASV by contact with infected rodents (*Mastomys natalensis*) or virus contaminated rodents excreta/secreta, while person-to-person transmission is possible, particularly in hospital environments. Clinically, LASV infections range from apparently asymptomatic or mild to severe hemorrhagic fever characterized by multi-organ failure, a condition referred as Lassa fever (J. McCormick and Fisher-Hoch, 2002). The most useful clinical predictors of Lassa fever are pharyngitis, fever, retrosternal pain, proteinuria for diagnosis, sore throat, and vomiting (Goeijenbier et al., 2013). Overall, LASV related death rates are about 2 %; however, in hospitalized cases, this incidence rises to 15 %, and in outbreak conditions, it can reach 50 % (McCormick et al., 1987). To prevent LASV infection or sickness in humans, there is currently no approved vaccine, despite the high prevalence and associated morbidity and mortality. Currently, supportive care and, where accessible, the early administration of ribavirin are the mainstays of treatment for Lassa fever patients.

Lassa virus infection was identified for the first time in 1969, in a Lassa town, Nigeria (Africa), where two medical missionaries death occurred and near-fatal illness of a third (Buckley et al., 1970; Frame, 1975). Subsequently, several West African countries have reported cases of the disease (Carey et al., 1972; Fraser et al., 1974; Monath et al., 1973). A large portion of the first epidemiologic data regarding Lassa

* Corresponding authors.
E-mail addresses: kattia1.c@ksu.edu.sa (K.A. Attia), lianghuiying@gdph.org.cn (H. Liang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.virusres.2025.199528
Received 29 July 2024; Received in revised form 17 November 2024; Accepted 17 January 2025
Available online 29 January 2025

fever came from studies conducted on hospitalized patients, many of whom had nosocomial infections. The majority of cases were linked to person-to-person viral transmission (Bowen et al., 1975; Carey et al., 1972; Frame, 1975; Frame et al., 1970; Fraser et al., 1974; Monath et al., 1973), which is not a characteristic of the other human *Arenavirus* infections that are currently known to exist.

Once the Lassa virus was isolated from *Mastomys natalensis* in 1972, it was discovered that the virus was an *Arenavirus*, indicating the presence of a rodent host (Monath et al., 1974). Although their roles in human disease are still unknown, the isolation of Lassa virus-related viruses from several rodent species from other parts of Africa (Central African Republic, Zimbabwe, Mozambique, and South Africa) has now increased our understanding of African *Arenavirus* epidemiology (Johnson et al., 1981; Wulff et al., 1977). In rural West Africa, the relative significance of rodent-to-person and person-to-person transmission in the natural history of Lassa fever has not been adequately determined. Designing epidemiologic studies that would enable the independent identification of either modality proved to be challenging. The one study that has been done so far has shown that person-to-person transmission happens in villages, but it is impossible to say accurately how much it is important (Keenlyside et al., 1983). As a result, there are two main themes in the epidemiology of Lassa fever: the seeming dominant rodent-to-human propagation and the equally significant person-to-person spread. The basic route of transmission would seem to be the spread to humans from a chronically infected rodent host, and the rodent is undoubtedly the major reservoir. Codon usage bias (CUB), a widespread phenomenon in the realm of molecular genetics, pertains to the favoring of specific synonymous codons over others during protein synthesis. Understanding the mechanisms driving CUB provides valuable information about gene evolution and regulation, as well as the optimization of protein expression and function (Kumar et al., 2016a; Nguyen et al., 2021; Quax et al., 2015). CUB is not limited to a particular group of organisms but is evident across diverse life forms, including bacteria, archaea, viruses, fungi, plants, and animals (Butt et al., 2016; Rahman et al., 2018). Therefore, studying codon usage patterns in various organisms can reveal unique adaptations and evolutionary processes that have shaped their genomes. CUB plays a crucial role in influencing the chemical variations among amino acids, as the frequency of different codons also affects the likelihood of errors occurring during protein synthesis (Tyagi et al., 2022; Zhang et al., 2011). Codon usage preferences vary among genes, even within the same genome (Yang et al., 2021). The three main factors that contribute to CUB are mutation pressure, natural selection, and random genetic drift (Deb et al., 2021, 2020). Additionally, several other factors play a role, such as length of gene and level of expression, protein structure, RNA stability, tRNA abundance, and GC content (Rahman et al., 2022b, 2022c, 2018; Sharma et al., 2020). CUB is understood to develop due to the interplay between translational selection and the impact of mutation pressure (Kumar et al., 2016a). Studying the CUB of viruses can yield valuable information about the control of viral gene expression, which, in turn, can contribute to the development of more effective vaccines (Butt et al., 2016; Rahman et al., 2018). Compared with prokaryotic and eukaryotic genomes, the viral genome has certain features such as dependence on its hosts for replication, synthesizing protein, and transmission of proteins. The interaction between virus and host is considered to affect survival, adaptation, and evolution of virus, and escape from hosts immune system (Burns et al., 2006; Costafreda et al., 2014; Mueller et al., 2006). Changes in codon usage pattern were shown to be important in major ecological shifts in virus evolution (Sun et al., 2020). The purpose of this work is to examine codon usage patterns by employing many codon bias indicators. This study is crucial for a better understanding of the structural organization and molecular evolution of LASV genome.

2. Material and methods

2.1. Collection of data

In the present study, the complete coding sequences of 162 Lassa virus (LASV) strains (listed in Table S1) were obtained from the NCBI GenBank database (https://ncbi.nlm.nih.gov/nuccore/?term=lassa+virus). The codon usage data of LASV hosts; *Homo sapiens* and *Mastomyous Natalensis* were downloaded from the codon usage database (https://www.kazusa.or.jp/codon/).

2.2. Analysis of nucleotide composition

The CodonW 1.4.4 program (http://codonw.sourceforge.net//culong.html) was used to compute several factors related to nucleotide composition and distribution. These include the total compositions of individual nucleotides (A%, C%, T%, and G%), the collective frequencies of nucleotides (AT% and GC%), the distribution of nucleotides at the 3rd position of codons (A3 %, G3 %, C3 %, and T3 %), the frequencies of G + C content at various positions within codons (GC1 s, GC2 s, GC12 s, and GC3 s), and the comprehensive composition percentages of AT%, AT3 %, and GC%. If there is no external pressure, mutations should occur randomly at any codon position; otherwise, they will occur in a particular direction (Behura and Severson, 2013). The preference for a certain base will vary at each of the three codon sites if selection pressure is present; otherwise, the base composition will be the same at each codon site (Wang et al., 2023b).

2.3. Analysis of relative synonymous codon usage (RSCU)

RSCU is a frequently employed parameter to analyze the overall variation in synonymous codon usage across genes (Sharp and Li, 1986). It is estimated by comparing how frequently each codon appears in protein-coding genes to the average usage of synonymous codons. The RSCU values for 162 LASV coding sequences were computed using the CAIcal (http://genomes.urv.es/CAIcal) (Puigbò et al., 2008). Codons having RSCU values exceeding 1.0 demonstrate a positive bias in their usage, indicating their abundance. Conversely, codons with RSCU values below 1.0 show a negative bias in their usage, indicating their lower abundance (Sharp and Li, 1986). Codons that possess an RSCU value exceeding 1.6 were classified as over-represented, indicating a higher occurrence. Conversely, codons having an RSCU value below 0.6 were classified as under-represented, indicating a lower occurrence (Deb et al., 2021).

2.4. Effective number of codons (ENC) and ENC plot analysis

Effective number of codons (ENC) analysis was used to measure the absolute CUB of the LASV coding sequences. ENC, which falls within the range of 20 to 61, is utilized to evaluate the extent of codon usage bias within a gene, without taking into account its size or the quantity of amino acids it encodes (Wright, 1990). When the value of ENC is 20, it indicates a significant bias in codon usage, where a single synonymous codon for a particular amino acid is utilized. Conversely, an ENC value of 61 indicates the absence of bias, suggesting that all the synonymous codons available for a particular amino acid are utilized in equal proportions. To assess the level of CUB in LASV, ENC was computed by CodonW 1.4.4. The ENC-plot, which involves plotting ENC values against GC3 s, is frequently employed to analyze the impact of G + C compositional constraints on CUB (Wright, 1990). If the expected ENC values are situated above or near the standard curve, it suggests that codon usage is primarily affected by G + C mutation bias. However, if the ENC values are situated beneath the standard curve, this suggests that selection pressure have influenced the CUB (Wong et al., 2010). The expected ENC values were determined utilizing the equation (Yao et al., 2020).

$$ENC^{expected} = 2 + s + \left(\frac{29}{s^2 + (1-s)^2}\right)$$

where s represents the given GC3 s values.

2.5. Analysis of neutrality plot

The neutrality plot analysis relies on creating a regression line by plotting GC3 s against GC12s. It signifies how the interplay of mutational pressure and natural selection influences the formation of CUB in viral genes (Sueoka, 1988). A regression line with a slope approaching 1 indicates that mutational pressure mainly drives the CUB, with limited influence from selection pressure. Conversely, regression curves nearing zero indicate that selection pressure significantly influences the shaping of CUB (Yao et al., 2020). Using Codon W software (version 1.4.4), the GC3 s and GC12 s values for the LASV strains were computed.

2.6. Parity rule 2 plot analysis

According to parity rule 2 (PR2), the mononucleotides $T = A$ and C = G in the coding sequences show that there is no bias in the mutation and selection. To assess the factors that influence on the codon usage pattern, the PR2 is plotted with GC bias at third codon position [G3 /(G3 +C3)] as abscissa against AT bias, at third codon position [A3 /(A3 +T3)] as ordinate (Nguyen et al., 2021), and the origin at (0.5, 0.5) where G = C and $A = T$ points lying have no bias with no affect towards natural selection and mutation pressure (Rahman et al., 2022a).

2.7. Codon adaptation index analysis

Codon adaptation index (CAI) analysis is employed to estimate the expression level of genes and determine how effectively viral genes have adjusted to their host organisms by comparing them to the reference host RSCU. CAI were calculated using an online tool, “CAIcal”. CAI values range from 0 to 1. Higher CAI values indicating the potential for highly expressed genes, and vice versa (Tao et al., 2009). Furthermore, the e-CAI (expected CAI) was analysed using the online tool “CAIcal”. The values of RSCU for the hosts genome were retrieved from the database of codon usage.

2.8. Correspondence analysis (COA)

Correspondence analysis (COA) is a multivariate statistical technique for examining significant patterns in codon usage variation across genes (Karniychuk, 2016). In present study, correlation analysis was employed to investigate the variations in codon usage patterns among LASV strains. To rationalize the impact of amino acid content on codon usage, the correlation analysis omitted three stop codons and two specific codons (AUG and UGG). COA was conducted using the CodonW (http://sourceforge.net/projects/codonw), applying the RSCU values as the basis for the analysis. Due to the fact that the first two axes typically account for a significant portion of the data’s variance compared to the other axes, only the first two axes were used to plot the LASV strains.

2.9. Phylogenetic analysis

A phylogenetic tree was created by utilizing the nucleotide sequences of LASV strains to investigate the evolutionary connections among LASV strains. The maximum likelihood approach was used to build a phylogenetic tree and the tree was subjected to a 1000-replicate bootstrap test, employing MEGA version 11 (Kumar et al., 2016b), and for further visualization using the Interactive tree of Life (iTOL).

3. Results

3.1. Nucleotide composition in Lassa virus strains

In current study, we assessed the codon usage pattern and composition of Lassa virus strains by analyzing their coding sequences. It’s worth noting that CUB can be strongly impacted by the entire nucleotide composition of the genome (Kattoor et al., 2015; Liu et al., 2017; Moratorio et al., 2013). A prior research suggested that nucleotide bias is a major factor in shaping the codon usage specific to viruses, which in turn restricts the influence of codon selection and translational control (Rahman et al., 2023). Here, the nucleotide composition of the 162 Lassa virus strains was examined to assess the possible influence of compositional constraints on CUB (Table 1). The Lassa virus strains were observed to exhibit a relatively greater prevalence of A and T nucleotides in comparison to G and C nucleotides (Table 1). Among the four nucleotides, the highest average percentage was observed for T (29.18 %) and A (28.16 %). On the other hand, C (22.91) and G (19.73 %) had the smallest average value . Important are especially nucleotides at the third position of codons (A3, T3, G3, C3) (Jenkins and Holmes, 2003; Nasrullah et al., 2015; Wong et al., 2010). Regarding the third position within synonymous codons, T3 % (29.21) and A3 (28.19) had the highest frequency, while C3 % (22.91) and G3 (19.67) exhibited the lowest frequency. The mean composition of AT was found to be 57.34 %, while GC accounted for 42.65 % on average. Furthermore, the average AT3 composition was 57.40 %, and the average GC3 composition was 42.59 % (Table 1). Our findings demonstrated that the average nucleotide compositions of A and T were notably elevated in comparison to G and C.

Moreover, GC content at all positions of codons (GC1, GC2, and GC3) also plays an important role in influencing overall codon usage preferences (Jenkins and Holmes, 2003; Nasrullah et al., 2015; Wong et al., 2010). The first codon site had a GC value of 42.83 %, the second codon site had a GC value of 42.53 %, and the third codon site had a GC value of 42.59 %. Here the results showed that the GC content at each codon position is a reliable indicator of base composition bias.

3.2. Relative synonymous codon usage (RSCU) of Lassa virus strains

The examination of RSCU values was performed to uncover patterns in the synonymous codon usage. The mean RSCU values were assessed for LASV strains (Table 2). Despite the fact that more than one codon can specify the majority of amino acids, it is thought that only a small subset of possible codons was actually used in highly expressed genes. Here, the most frequently preferred codons ((Phe) TTT, (Leu) CTT, (Ile) ATT, Val GTT, (Ser) TCA, (Pro) CCA, (Thr) ACA, (Ala) GCA, (Tyr) TAT, His (CAT), Gln (CAA), (Asn) AAT, (Lys) AAA, (Asp) GAT, (Glu) GAA, (Cys) TGT, (Arg) AGA, (Gly) GGA) were found to be A/T-terminated (A-ended: 09; T-ended: 09). The RSCU analysis clearly indicated that the LASV strains showed a stronger codon usage bias against codons that terminate with A/T. The analysis of codons that were over-represented (RSCU $\geq$ 1.6) and under-represented (RSCU $\leq$ 0.6) revealed that the RSCU values of LASV strains ranged between 0.6 and 1.6. It was intriguing to observe that the majority of the underrepresented codons had C/G-endings (Table 2). Additionally, the RSCU for each separate LASV genomic segment shows the same patterns AT rich (Fig. S1).

Moreover, a comparison was made among the LASV and hosts (*Homo sapiens* and *Mastomyous Natalensis*) for the RSCU values of 59 sense codons (Fig. S2). Although there were some variations in the RSCU values of the 59 sense codons among the LASVs and hosts (*Homo sapiens* and *Mastomyous Natalensis*). But the general pattern remained quite consistent. We observed some over-represented and under-represented codons among LASV and *Mastomyous Natalensis*. In particular, Arg (AGG) is over-represented in both LASVs and *Mastomyous Natalensis*.

Table 1
Nucleotide compositional analysis of Lassa virus (LASV) strains.

Accession No	%A	%C	%T	%G	%A3	%C3	%T3	%G3	%GC	%AT	GC3 %	AT3 %	ENC
KM822102	28.71	22.83	30.21	18.25	31.4	16.69	32.6	19.3	41.08	58.92	36	64	52
KM822007	29.18	22.52	30.51	17.78	31.27	16.85	32.52	19.36	40.3	59.7	36.21	63.79	52
KM822131	29.08	23.65	30.12	17.15	24.47	28.88	30.74	15.9	40.8	59.2	44.79	55.21	50.7
KM822080	27.6	23.1	28.05	21.25	26.59	21.67	29.54	22.2	44.34	55.66	43.87	56.13	55
KM821889	26.88	23.56	28.09	21.47	28.53	22.7	27.47	21.29	45.02	54.98	43.99	56.01	54.1
KM822128	27.9	22.24	28.72	21.13	29.07	21.8	28.37	20.75	43.37	56.63	42.56	57.44	53.5
KM821991	29.21	22.55	30.54	17.7	31.45	16.8	32.59	19.16	40.25	59.75	35.96	64.04	52.4
KM821881	29.05	23.58	30.06	17.31	31.54	17.24	32.91	18.32	40.89	59.11	35.56	64.44	51.3
KM822040	28.93	22.75	30.28	18.04	23.36	28.23	31.32	17.09	40.79	59.21	45.32	54.68	50.4
KM822095	27.58	22.57	28.59	21.26	30.78	20.82	28.38	20.02	43.83	56.17	40.84	59.16	52
KM822073	28.98	22.62	30.37	18.03	31.93	22.25	27.87	17.95	40.65	59.35	40.2	59.8	51.6
KM821909	26.78	23.39	28.32	21.5	28.58	22.57	27.79	21.06	44.9	55.1	43.63	56.37	53.9
KM822070	29.03	22.62	30.37	17.97	23.74	28.19	31.22	16.85	40.6	59.4	45.04	54.96	50.8
KM822082	27.51	23.07	28.21	21.21	26.48	22.42	29.21	21.89	44.28	55.72	44.31	55.69	54.1
KM822126	27.08	23.44	28.27	21.21	26.13	26.49	26.76	20.62	44.65	55.35	47.11	52.89	53
KM822077	28.97	22.56	30.47	17.99	23.43	28.55	31.18	16.84	40.56	59.44	45.39	54.61	50.6
KM822036	28.91	22.4	30.62	18.08	31.57	16.63	32.56	19.24	40.48	59.52	35.87	64.13	51.8
KM822110	28.74	23.66	30.11	17.49	31.65	17.46	32.77	18.12	41.15	58.85	35.58	64.42	52
KM822056	28.88	22.79	30.26	18.07	31.72	23.24	27.05	17.99	40.86	59.14	41.23	58.77	51.3
KM822115	26.55	23.46	28.38	21.6	25.82	26.7	26.61	20.87	45.06	54.94	47.57	52.43	54.3
KM822103	27.7	22.78	28.32	21.19	27.14	21.31	30.5	21.04	43.97	56.03	42.35	57.65	53.4
KM821904	27.72	23.06	28.73	20.49	30.2	21.95	28.49	19.35	43.55	56.45	41.31	58.69	53.7
KM822062	27.83	22.85	28.12	21.2	26.94	21.02	30.3	21.73	44.05	55.95	42.76	57.24	54.4
KM822071	27.87	23.05	28.04	21.04	25.89	25.98	26.77	21.37	44.09	55.91	47.34	52.66	52.8
KM822000	27.63	22.78	28.4	21.19	27.07	21.78	29.63	21.52	43.97	56.03	43.3	56.7	52.8
KM821888	28.84	23.76	30.05	17.35	31.15	25.18	26.49	17.18	41.11	58.89	42.36	57.64	51.6
KM821901	28.62	23.26	30.54	17.58	29.88	23.88	27.89	18.35	40.84	59.16	42.23	57.77	52.1
KM822085	29.17	22.72	30.33	17.78	31.29	17.26	32.2	19.25	40.51	59.49	36.51	63.49	52.6
KM822129	29.06	23.59	30.19	17.16	24.42	28.83	30.77	15.98	40.75	59.25	44.81	55.19	50.8
KM821990	27.63	22.55	28.54	21.28	30.34	20.9	28.57	20.19	43.83	56.17	41.09	58.91	51.9
KM822008	27.49	22.85	28.34	21.32	26.48	21.7	29.85	21.97	44.17	55.83	43.67	56.33	53.7
KM822061	29.04	22.65	30.4	17.91	31.8	16.54	32.63	19.03	40.56	59.44	35.57	64.43	51.3
KM822032	28.57	22.81	30.22	18.4	31.36	23.18	26.98	18.47	41.21	58.79	41.65	58.35	51.3
KM822104	28.95	22.73	30.34	17.98	31.2	17.13	32.35	19.32	40.71	59.29	36.45	63.55	52.6
KM822109	27.52	22.92	28.26	21.3	26.73	21.59	29.73	21.95	44.22	55.78	43.54	56.46	54.9
KM822050	29.1	22.72	30.25	17.93	23.94	28.48	31	16.58	40.65	59.35	45.06	54.94	50.2
KM821893	28.61	23.2	30.52	17.67	29.99	23.82	27.71	18.48	40.87	59.13	42.29	57.71	51.8
KM822059	27.6	22.98	28.1	21.33	29.99	21.83	27.6	20.59	44.31	55.69	42.41	57.59	52.2
KM822031	27.58	22.92	28.14	21.36	26.94	21.21	29.84	22.01	44.28	55.72	43.22	56.78	54.3
KM821999	28.59	22.71	30.33	18.37	23.33	28.51	31.11	17.05	41.08	58.92	45.55	54.45	50.2
KM822112	28.96	23.73	29.89	17.42	24.08	29.43	30.47	16.02	41.15	58.85	45.45	54.55	50.3
KM822084	27.62	23	28	21.37	26.64	21.14	30.02	22.2	44.38	55.62	43.34	56.66	54.1
KM822038	28.98	22.63	30.37	18.02	23.6	28.61	30.96	16.83	40.65	59.35	45.43	54.57	50.9
KM822041	27.32	22.83	28.35	21.51	26.5	21.83	29.67	22.01	44.34	55.66	43.84	56.16	54
KM821993	29.13	22.26	30.7	17.91	23.65	28.17	31.51	16.67	40.17	59.83	44.84	55.16	50.5
KM822099	27.87	22.73	28.4	20.99	30.65	21.35	28.08	19.93	43.73	56.27	41.28	58.72	52.6
KM822069	27.28	22.67	28.58	21.47	29.78	20.53	29.16	20.53	44.14	55.86	41.06	58.94	51.7
KM821903	28.91	23.38	30.41	17.3	31.52	24.37	27.02	17.09	40.69	59.31	41.46	58.54	50.6
KM822079	29.44	22.49	30.5	17.56	32.86	22.77	27.14	17.23	40.06	59.94	40	60	51.2
KM822017	29.1	22.66	30.33	17.9	32.05	22.68	27.53	17.74	40.56	59.44	40.42	59.58	50.7
KM821898	29.11	23.74	29.89	17.26	24.32	28.83	30.65	16.21	41	59	45.04	54.96	50.5
KM821996	27.72	22.22	28.87	21.19	30.51	20.55	28.75	20.19	43.42	56.58	40.74	59.26	51.1
KM822087	28.65	22.64	30.46	18.25	31.21	23.29	27.1	18.4	40.89	59.11	41.69	58.31	51.2
KM822047	27.59	22.66	28.32	21.43	30.52	21.28	28.06	20.14	44.09	55.91	41.42	58.58	51.9
KM822022	27.58	23.09	27.96	21.36	26.76	21.74	29.4	22.1	44.45	55.55	43.84	56.16	54.8
KM821897	27.35	23.25	28.41	20.98	29.35	22.37	28.03	20.25	44.24	55.76	42.62	57.38	54.6
KM821998	27.68	23.35	27.86	21.11	30.74	22.26	27.3	19.7	44.46	55.54	41.96	58.04	52.8
KM822018	27.92	23.41	27.77	20.91	26.59	22.97	28.45	22	44.32	55.68	44.96	55.04	54.1
KM821906	27.35	23.22	28.44	21	29.38	22.3	28.05	20.27	44.22	55.78	42.57	57.43	54.5
KM822027	27.03	22.69	28.53	21.75	25.77	21.2	30.26	22.78	44.44	55.56	43.98	56.02	54.6
KM822033	27.22	22.92	28.22	21.64	25.53	25.26	27.28	21.93	44.56	55.44	47.19	52.81	53
KM821995	29.11	22.75	30.3	17.84	24.01	28.27	31.37	16.35	40.59	59.41	44.62	55.38	50.4
KM821902	27.19	23.22	28.42	21.16	25.84	26.01	26.9	21.25	44.39	55.61	47.27	52.73	54.6
KM822111	27.16	23.12	28.55	21.16	30.99	23.42	23.6	21.99	44.29	55.71	45.41	54.59	52
KM822025	27.15	23.22	28	21.64	29.38	21.72	27.88	21.02	44.85	55.15	42.74	57.26	51.6
KM821883	28.87	23.75	30.05	17.33	24.12	28.66	30.89	16.33	41.09	58.91	44.99	55.01	50.8
KM822013	29.03	22.56	30.38	18.03	31.62	16.83	32.45	19.11	40.59	59.41	35.94	64.06	52.2
KM822101	27.42	23.11	28.04	21.42	25.89	26.15	26.51	21.45	44.53	55.47	47.61	52.39	53
KM822074	27.72	22.7	28.37	21.2	26.08	25.46	27.05	21.41	43.91	56.09	46.87	53.13	52.3
KM822028	28.91	22.93	30.15	18	23.55	28.68	30.99	16.78	40.94	59.06	45.45	54.55	50.7
KM821891	29.05	23.51	30.15	17.28	31.66	24.59	26.51	17.24	40.8	59.2	41.84	58.16	51.3
KM822086	27.9	23.1	28.13	20.87	26.98	21.69	29.89	21.43	43.97	56.03	43.12	56.88	54.5
KM822068	29.33	22.86	30.04	17.77	32.75	23.37	26.48	17.39	40.63	59.37	40.76	59.24	51
KM822042	28.75	22.5	30.42	18.33	31.47	17.68	32.55	18.3	40.83	59.17	35.98	64.02	52.8

(*continued on next page*)

Table 1 (*continued*)

Accession No	%A	%C	%T	%G	%A3	%C3	%T3	%G3	%GC	%AT	GC3 %	AT3 %	ENC
KM822106	29.23	22.66	30.5	17.61	31.57	16.78	32.52	19.14	40.27	59.73	35.92	64.08	52.3
KM822116	29.25	23.24	30.36	17.15	31.5	25.06	26.55	16.89	40.39	59.61	41.96	58.04	51.4
KM821992	27.45	23.1	28.13	21.32	30.24	21.23	28.28	20.25	44.42	55.58	41.48	58.52	52
KM822083	28.93	22.82	30.26	17.99	23.54	28.71	31.04	16.71	40.81	59.19	45.42	54.58	50.6
KM822009	29.11	22.41	30.6	17.88	23.53	28.28	31.4	16.78	40.29	59.71	45.06	54.94	50.5
KM822088	27.41	22.56	28.65	21.38	26.94	21.29	30.21	21.55	43.93	56.07	42.84	57.16	52.1
KM822097	27.74	23.03	28.04	21.18	26.84	21.12	29.83	22.21	44.21	55.79	43.34	56.66	54.5
KM821994	27.53	22.99	28.09	21.39	29.87	21.78	28	20.36	44.39	55.61	42.13	57.87	52.1
KM821884	26.88	23.56	28.09	21.47	28.53	22.7	27.47	21.29	45.02	54.98	43.99	56.01	54.1
KM822035	27.53	22.81	28.38	21.28	26.03	25.51	26.82	21.64	44.09	55.91	47.14	52.86	51.9
KM822024	29.2	22.84	30.23	17.73	31.4	17.07	32.27	19.26	40.56	59.44	36.32	63.68	51.9
KM15076	29.05	21.26	26.99	22.71	30.8	19.02	30.07	20.11	43.96	56.04	39.13	60.87	54.1
KM822064	29.24	22.48	30.66	17.62	32.68	22.36	27.73	17.24	40.1	59.9	39.6	60.4	50.1
KM822001	29.14	22.86	30.2	17.8	32.37	23.39	27.03	17.22	40.66	59.34	40.6	59.4	51.3
KM822120	26.97	23.35	28.39	21.29	29.06	23.32	26.94	20.67	44.64	55.36	43.99	56.01	53.8
KM822046	29.17	22.75	30.33	17.74	31.36	16.9	32.44	19.3	40.5	59.5	36.2	63.8	51.8
KM822054	28.67	22.77	30.25	18.31	30.98	17.31	32.26	19.45	41.08	58.92	36.76	63.24	52.4
KM822055	27.32	23.05	28.15	21.49	25.55	25.38	27.32	21.75	44.53	55.47	47.13	52.87	52.9
KM821892	26.73	23.36	28.44	21.47	28.94	22.39	27.79	20.88	44.84	55.16	43.27	56.73	54.6
KM822026	28.65	22.8	30.35	18.19	23.49	28.45	31.26	16.8	41	59	45.25	54.75	50.6
KM822122	26.96	23.3	28.32	21.42	25.49	25.66	27.26	21.59	44.72	55.28	47.26	52.74	53.8
KM821905	28.62	23.17	30.53	17.69	31.41	17.26	32.99	18.34	40.86	59.14	35.6	64.4	52.4
KM822081	29.17	22.91	30.21	17.71	32.84	22.83	27.11	17.23	40.63	59.37	40.06	59.94	51
KM822002	27.62	22.94	28.15	21.29	27.03	21.91	29.33	21.73	44.23	55.77	43.64	56.36	54.1
KM822075	28.97	22.59	30.5	17.95	31.61	16.73	32.7	18.95	40.53	59.47	35.68	64.32	52
KM822127	29.04	24.12	29.58	17.26	30.16	25.12	26.73	17.99	41.38	58.62	43.11	56.89	51.1
KM822107	27.43	22.75	28.42	21.4	26.72	21.87	29.63	21.78	44.15	55.85	43.65	56.35	53.9
KM822114	28.96	23.82	30.02	17.2	24.25	28.66	31.13	15.96	41.02	58.98	44.62	55.38	51
KM822098	28.55	22.87	30.21	18.36	31.35	17.01	32.43	19.21	41.23	58.77	36.22	63.78	52.4
KM822004	27.3	22.81	28.43	21.45	26.6	21.72	29.79	21.9	44.27	55.73	43.62	56.38	53.2
KM822132	26.67	23.31	28.53	21.49	28.74	22.19	28.12	20.95	44.8	55.2	43.15	56.85	54.5
KM822057	27.72	23.11	27.84	21.32	26.87	21.32	29.87	21.94	44.43	55.57	43.26	56.74	53.8
KM822130	26.68	23.38	28.47	21.47	26.06	26.24	26.94	20.76	44.85	55.15	47	53	54.9
KM822010	27.6	22.94	28.05	21.42	25.94	25.85	26.65	21.57	44.36	55.64	47.42	52.58	52.9
KM822094	29.05	22.87	30.26	17.81	31.97	23.34	26.93	17.76	40.68	59.32	41.1	58.9	52
KM822093	27.45	22.99	28.25	21.32	25.6	25.42	27.3	21.68	44.31	55.69	47.1	52.9	52.6
KM822015	29.09	22.7	30.32	17.89	23.56	28.33	31.44	16.67	40.59	59.41	45	55	50.4
KM822092	29.21	22.57	30.48	17.75	31.2	17.06	32.41	19.32	40.32	59.68	36.39	63.61	52
KM822108	28.97	22.67	30.35	18.01	31.42	17	32.38	19.2	40.68	59.32	36.2	63.8	52.2
KM821882	26.9	23.16	28.54	21.4	28.66	22.31	28.13	20.9	44.56	55.44	43.21	56.79	53.7
KM822034	28.73	22.86	30.26	18.15	31.6	23.38	27.14	17.89	41	59	41.26	58.74	51.4
KM822096	29.07	22.54	30.53	17.86	32.2	22.77	27.19	17.84	40.4	59.6	40.61	59.39	51.4
KM822049	27.69	23.05	28.16	21.09	26.43	21.85	29.6	22.11	44.14	55.86	43.96	56.04	54.1
KM822053	27.58	22.91	28.14	21.38	20.62	27.84	30.93	20.62	44.29	55.71	48.46	51.54	53.2
KM822090	29.14	22.85	30.22	17.8	23.59	28.44	31.29	16.68	40.65	59.35	45.12	54.88	49.8
KM822039	27.61	22.97	28.16	21.26	29.87	21.5	28.19	20.44	44.23	55.77	41.94	58.06	51.2
KM03362	30.82	19.89	24.97	24.32	27.87	21.15	27.7	23.28	44.21	55.79	44.43	55.57	51.9
KM822048	28.81	22.49	30.51	18.19	31.5	17.11	32.33	19.06	40.68	59.32	36.17	63.83	52.6
KM822124	26.9	23.37	28.31	21.43	28.92	23.1	26.98	20.99	44.8	55.2	44.09	55.91	54
KM822078	27.32	22.99	28.23	21.46	26.53	21.66	29.71	22.1	44.44	55.56	43.77	56.23	55.3
KM821997	28.83	22.7	30.31	18.16	31.19	23.12	26.97	18.72	40.86	59.14	41.84	58.16	52.2
KM822014	27.75	22.93	28.28	21.04	26.86	21.63	29.79	21.72	43.97	56.03	43.35	56.65	53.1
KM822044	28.84	22.95	30.11	18.09	31.68	16.89	32.42	19	41.04	58.96	35.89	64.11	52.4
KM821908	28.49	23.93	29.71	17.87	23.97	29.45	30.23	16.35	41.8	58.2	45.8	54.2	50.1
KM822117	26.86	23.37	28.4	21.37	25.98	26.42	26.77	20.83	44.74	55.26	47.25	52.75	54.5
KM822091	26.94	22.95	28.29	21.83	25.7	21.65	29.49	23.15	44.78	55.22	44.81	55.19	53.4
KM822052	29	22.89	30.12	17.98	23.69	27.84	31.78	16.68	40.87	59.13	44.52	55.48	49.9
KM822021	29.24	22.66	30.37	17.73	23.57	28.16	31.56	16.71	40.39	59.61	44.87	55.13	50.5
KM822005	29.18	22.25	30.64	17.92	23.53	28.15	31.28	17.03	40.18	59.82	45.19	54.81	50.9
KM822020	27.68	23	28.06	21.26	30.04	21.47	28.09	20.41	44.26	55.74	41.87	58.13	52.8
KM822076	27.48	22.81	28.4	21.32	25.78	25.96	26.76	21.5	44.13	55.87	47.46	52.54	53.1
KM822011	28.89	22.57	30.44	18.09	31.45	22.49	27.63	18.42	40.66	59.34	40.91	59.09	51.6
KM822043	27.51	22.75	28.37	21.37	26.63	22.05	29.28	22.05	44.12	55.88	44.09	55.91	53.9
KM822051	27.67	23.2	27.93	21.2	27.14	21.5	29.87	21.5	44.41	55.59	43	57	54.6
KM822045	27.43	22.64	28.49	21.44	29.25	21.06	28.55	21.15	44.08	55.92	42.2	57.8	52.3
KM822029	27.46	22.98	28.08	21.48	30.02	21.74	27.73	20.51	44.45	55.55	42.25	57.75	52
KM821896	28.61	23.14	30.55	17.7	30.34	23.37	27.73	18.56	40.83	59.17	41.93	58.07	52
KM822012	27.56	22.93	28.23	21.28	26.88	22.02	29.44	21.66	44.21	55.79	43.68	56.32	53.2
KM822100	29.02	22.38	30.64	17.95	31.64	16.55	32.6	19.21	40.33	59.67	35.76	64.24	51.5
KM822006	27.94	22.99	28.18	20.9	26.97	21.93	29.44	21.66	43.88	56.12	43.59	56.41	54
KM822113	27.07	23.18	28.41	21.34	30.72	23.69	22.89	22.71	44.52	55.48	46.39	53.61	52
KM822123	28.88	23.36	30.25	17.52	23.89	28.5	31.27	16.35	40.88	59.12	44.84	55.16	49.9
KM822003	29.21	22.48	30.69	17.62	31.77	17.93	30.3	20	40.1	59.9	37.93	62.07	51.5
KM821899	26.71	23.35	28.53	21.41	28.45	22	28.62	20.94	44.76	55.24	42.93	57.07	54
KM822065	27.65	23.14	27.91	21.31	26.08	25.91	26.61	21.4	44.44	55.56	47.3	52.7	52.8
KM821894	27.3	23.27	28.42	21.01	29.3	22.42	28.07	20.21	44.28	55.72	42.63	57.37	54.7

(*continued on next page*)

Table 1 (*continued*)

Accession No	%A	%C	%T	%G	%A3	%C3	%T3	%G3	%GC	%AT	GC3 %	AT3 %	ENC
KM822019	29.18	22.64	30.36	17.82	23.45	28.44	31.22	16.89	40.46	59.54	45.33	54.67	50.6
KM822016	27.63	23.05	28.19	21.13	31.13	21.34	28.22	19.31	44.18	55.82	40.65	59.35	52
KM822037	27.37	22.94	28.16	21.53	29.52	21.76	27.93	20.79	44.46	55.54	42.56	57.44	52.3
J04324	29.01	21.4	26.84	22.75	31.48	22.49	25.57	20.46	44.15	55.85	42.95	57.05	53.3
KM822058	29.13	22.59	30.58	17.7	31.67	23.03	27.39	17.91	40.29	59.71	40.94	59.06	51.2
KM822030	28.98	22.56	30.43	18.03	31.24	17.19	32.31	19.26	40.59	59.41	36.45	63.55	52.2
KM822105	27.45	23.04	28.1	21.42	26.13	25.68	26.92	21.27	44.45	55.55	46.95	53.05	53
AF246121	27.05	23.36	28.61	20.97	25.84	25.4	27.79	20.97	44.34	55.66	46.37	53.63	55
AY628201	28.89	21.04	27.13	22.95	24.01	22.42	30.27	23.3	43.98	56.02	45.72	54.28	53.2
GU830839	26.81	23	28.98	21.21	25.29	26.18	27.35	21.18	44.21	55.79	47.36	52.64	53.9
KF478765	26.39	24.37	28.18	21.06	25.21	24.37	28.82	21.6	45.43	54.57	45.97	54.03	54.2
KF478766	26.71	23.36	28.31	21.62	25.53	22.07	28.98	23.42	44.99	55.01	45.49	54.51	53.5
Mean	28.16	22.91	29.18	19.73	28.19	22.91	29.21	19.67	42.65	57.34	42.59	57.40	52.42
STD	0.87	0.50	1.12	1.83	2.99	3.54	2.10	2.00	1.84	1.84	3.39	3.39	1.39

NC represents the effective number of codons.
GC1 represents the G + C content at the first position of codons.
GC2 represents the G + C content at the second position of codons.
GC3 represents the G + C content at the third positions of codons.
AU3 represents the *A* + U content at the third positions of codons.

3.3. Effective number of codon (ENC) and ENC plot analysis

A gene's bias caused by an uneven codon usage is measured through ENC. ENC values were computed to measure the degree of CUB within the coding region of LASV (Table 1). Here, the ENC values for LASV strains ranged between 49.8 and 55.3, with an average value of 52.42. The ENC values of all the LASV strains were lower than the expected values, suggesting that mutational pressure play role in determining the codon usage patterns within these genes. The analysis has revealed that with mean ENC values>35, the CUB and gene expression of LASV strains were both low and exhibited a slight bias.

Next, the ENC values were graphed against the GC3 s contents to ascertain the principal factor that affects the CUB in LASV strains. As depicted in Fig. 1, all points are located below and near the expected curve. This suggests that, beyond mutation bias, selection exerted an impact on the development of codon bias in the LASV. Nonetheless, using ENC-plot analysis by itself is insufficient for precise evaluation and differentiation of the impacts of mutation pressure and selection pressure.

3.4. Neutrality plot analysis

The neutrality plot analysis shows that no significant correlation was found between GC3 and GC12 contents because the regression value and link are $P > 0.05$ and $r = 0.01$ (Fig. S3). This finding suggests that both natural selection and mutational pressure have an impact on the codon usage shaping of LASV.

3.5. Parity plot (PR2) analysis

The PR2 plot was plotted in order to investigate the factors influencing the codon usage bias in the LASV. The AT bias was framed on y-axis and GC on x-axis. Here, the analysis revealed an unequal usage of codons at the third codon base position between AT and GC (Fig. 2). This suggests that, in addition to the mutation, natural selection may have influenced the patterns of codon usage in LASV.

3.6. Codon adaptation index analysis

CAI values were determined for every codon found in the LASV strains, with reference to the codon usage data of host organisms (*Homo sapiens* and *Mastomyous Natalensis*). Here, the CAI analysis unveiled that the LASV strains (> 0.5), have notably adapted to their host organisms, which allowed it to use the translation source of hosts more efficiently. To check the observed significant statistical differences arises in the values of CAI (Puigbò et al., 2008; Rahman et al., 2018), the values of expected CAI (e-CAI) were considered for LASV CDS with hosts (*Homo sapiens* and *Mastomyous Natalensis*) codon usage sets. The result of the e-CAI value in relation to *Homo sapiens* and *Mastomyous Natalensis* revealed that the generated sequences keep to a normal distribution.

3.7. Correspondence analysis

Correspondence analysis (COA) was employed for the analysis of codon usage data in LASV strains. Based on the RSCU, the correspondence analysis was carried out using LASV strains. According to the results, axis 1 explained 25.71 % of the total variations, while axis 2 contributed to 15.82 % of the overall variations (Fig. 3). These findings imply that the first axis corresponds to the virus strain, and the second axis to the countries where the infection originates. Different geographical lineages are represented in the major axis's scattered data along with their connections to one another. The COA revealed that LASV strains were distributed and that there was no evidence of clustering with each other (Fig. 3).

3.8. Phylogenetic analysis

The maximum likelihood method was used for phylogenetic analysis in order to assess the impact of evolutionary relationship on the LASV codon usage pattern. Similar to the COA pattern (Fig. 3), the phylogenetic analysis suggested that the LASV strains are circulated all over the world, as none of the strains were involved to make cluster among various individual countries (Fig. S4).

4. Discussion

Here, we check out the use of synonymous codons in coding sequences from 162 LASV genomes to understand how the virus evolved molecularly under the influence of host, viral, and environmental factors. Prior studies have shown that the codon usage bias, or the preference for one type of codon over another, can be strongly influenced by the overall genetic composition (Jenkins, 2003). An analysis of the nucleotide composition of LASV genomes revealed that A/T nucleotides constitute the majority of the total nucleotide composition. This result was consistent with previous studies, i.e., Equine influenza virus (Bera et al., 2017), SARS-COV-2 (Datta et al., 2020), and Rabies virus (Li et al., 2023), wherein A/T frequencies were higher than G/C frequencies (Rahman et al., 2022). This finding also supports the prior studies (van Hemert and Berkhout, 2016; Rahman et al., 2018). Although the scientific importance for increased A, and decreased C is still unknown, it is

Table 2
The relative synonymous codon usage (RSCU) patterns of Lassa virus (LASV), and its hosts (*Homo sapiens* and *Mastomyous Natalensis*).

AA	Codons	LASV	*Homo sapiens*	*Mastomyous Natalensis*	AA	Codons	LASV	*Homo sapiens*	*Mastomyous Natalensis*
Phe	**TTT**	1.13	0.97	0.51	**Ala**	GCT	1.31	1.01	1.29
	TTC	0.87	1.03	1.49		GCC	0.94	1.12	**1.82**
Leu	TTA	0.82	**0.5**	0.1		**GCA**	1.53	1.18	0.79
	TTG	1.26	1 .0	1.25		GCG	**0.23**	**0.07**	**0.11**
	CTT	1.34	0.81	0.62	**Tyr**	**TAT**	1.22	0.71	0.87
	CTC	0.97	1.07	0.87		TAC	0.78	1.29	1.13
	CTA	**0.57**	**0.46**	1.57	**His**	**CAT**	1.15	0.85	1.34
	CTG	1.05	**2.33**	**1.6**		CAC	0.85	1.15	0.66
Ile	**ATT**	1.12	1.13	0.99	**Gln**	**CAA**	1.17	**0.49**	0.67
	ATC	1.05	1.37	**1.8**		CAG	0.83	1.51	1.33
	ATA	0.83	**0.5**	**0.22**	**Asn**	**AAT**	1.06	0.98	0.85
Val	**GTT**	1.23	0.79	**0.49**		AAC	0.94	1.02	1.15
	GTC	1.09	0.9	**0.56**	**Lys**	**AAA**	1.06	0.88	1.01
	GTA	0.74	**0.52**	1.34		AAG	0.94	1.12	0.99
	GTG	0.95	**1.79**	**1.61**	**Asp**	**GAT**	1.1	0.99	1.16
Ser	TCT	1.47	1.15	**2.18**		GAC	0.9	1.01	0.84
	TCC	1.0	1.17	1.21	**Glu**	**GAA**	1.09	0.85	**0.56**
	TCA	1.49	0.93	1.06		GAG	0.91	1.15	1.44
	TCG	**0.18**	**0.36**	**0.12**	**Cys**	**TGT**	1.18	0.95	1.36
	AGT	0.97	0.98	0.68		TGC	0.82	1.05	0.64
	AGC	0.89	1.42	0.76	**Arg**	CGT	**0.38**	**0.54**	0.75
Pro	CCT	1.32	1.2	1.3		CGC	**0.35**	1.11	1.08
	CCC	0.97	1.22	1.45		CGA	**0.49**	0.76	**0.57**
	CCA	1.43	1.14	1.05		CGG	**0.39**	1.31	**0.57**
	CCG	**0.28**	**0.45**	**0.2**		**AGA**	**2.58**	1.18	1.03
Thr	ACT	1.28	1.03	1.44		AGG	**1.81**	1.01	**1.74**
	ACC	1.0	1.32	1.34	**Gly**	GGT	1.05	0.71	0.81
	ACA	1.55	1.19	0.97		GGC	0.75	1.35	1.53
	ACG	**0.16**	**0.46**	**0.25**		**GGA**	1.11	1.01	0.97
						GGG	1.09	0.93	0.69

AA represents amino acid; the "RSCU" value represents the pattern of relative synonymous codon usage; over-represented (RSCU > 1.6), and under-represented (RSCU < 0.6) codons are marked with Red and orange, respectively; and the preferred codons for Lassa virus are Bold and Underlined.

essential to know the causes of these drifts in virus-related RNA genomes (van Hemert and Berkhout, 2016).

The RSCU analysis also showed that the LASV genome has a stronger codon usage bias toward A and T-ended codons (Table 2). These findings suggest that the primary driver of codon usage patterns in the LASV strains is mutational pressure, which is in line with prior studies (Ma et al., 2015; Wang et al., 2023a). Based on the analysis of both nucleotide content and RSCU, we infer that the preference for specific codons is largely influenced by compositional constraints, which in turn are determined by the presence of mutational pressure. The widely accepted viewpoint is that variations in nucleotide composition among different species are primarily driven by directional mutation pressure, a phenomenon predominantly influenced by the bias in AT/GC content (Sueoka, 1992). Several poxviruses, iridoviruses, and the African swine fever virus exhibit low GC contents. According to the mutation pressure hypothesis, this could be attributed to their replication occurring in the cytoplasm, especially if the nucleotide composition in this environment significantly contrasts with that of the cell nucleus (Dixon et al., 2013). Therefore, once it is proved that there is codon bias toward A and T-ended codons in LASV genomes, we next determined the patterns of codon usage and selection of the preferred codons in LASV genomes showed synergism to those of the hosts, *Homo sapiens* and *Mastomyous Natalensis* (Fig. S2). These findings align with the outcomes observed in the analysis of codon bias within RNA viruses like the Equine influenza viruses (EIVs) (Kumar et al., 2016a).

Next in LASV coding sequences, the mean ENC value was found to be 52.42, indicating slightly biased, comparatively stable, and conserved genomic composition among various LASV genomes. Previous studies

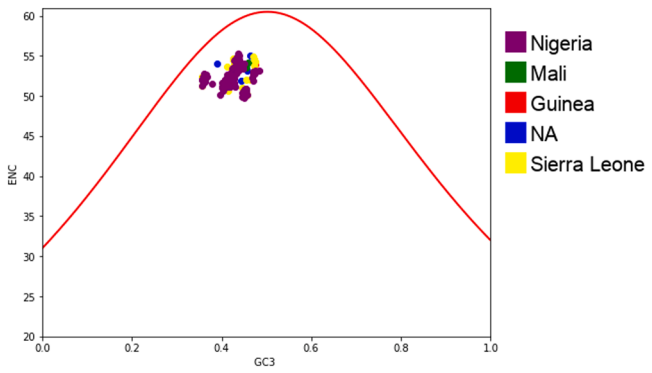

Fig. 1. ENC-GC3 plots of 162 Lassa virus (LASV) strains. The ENC (ENC-values, Y-axis) is plotted against the GC3 (GC3-values, X-axis). The dots in different color represent the different countries strain from where they originate.

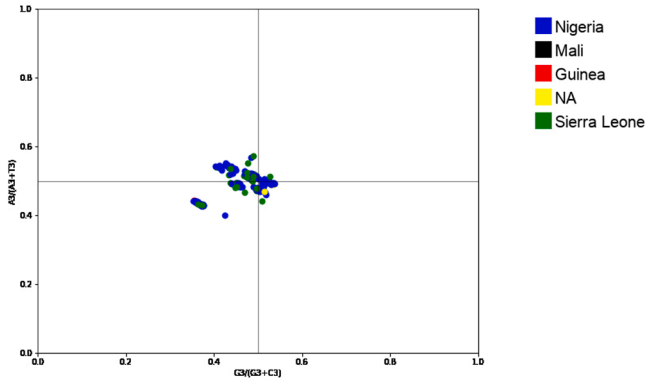

Fig. 2. Parity rule 2 (PR2) plot [A3/(A3+U3) against G3 /(G3+ C3)]. PR2 plot was calculated for 162 strains of LASV.

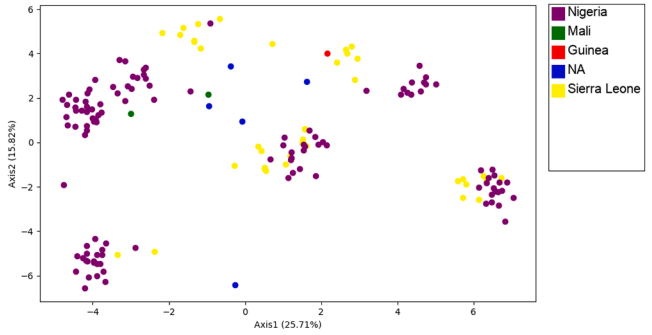

Fig. 3. Correspondence analysis (COA) of the genes in Lassa virus (LASV) genomes. The values of the first axis account for 25.71 % and the second axis account for 15.82 % of the total variation of each poly-protein coding region of LASV in COA.

have demonstrated that there exists an inverse relationship between ENC values and gene expression (Wright, 1990). Our results were aligned with the outcomes observed in the analysis of codon bias within RNA viruses like the Alongshan virus (Rahman et al., 2022), Crimean-Congo hemorrhagic fever virus (Rahman et al., 2018) and ZIKV (ENC, 53.21) (Butt et al., 2014; Cristina et al., 2016; Rahman et al., 2022a, 2018; Yao et al., 2020). It appears that LASV survival within the host has been made possible by the development of a low codon bias within LASV coding sequences, even though the host has codon usage preferences different from those of LASV. Following that, an ENC plot is commonly used to determine variation in codon usage among genes in different organisms (Butt at al., 2014). When the LASV ENC and GC3 values were plotted, it was discovered that all the isolates clustered tightly below the curve, demonstrating the significant influence of mutation pressure on LASV codon usage patterns as well as the influence of natural selection to some extent. Additionally, it has been previously reported that both natural selection and mutation pressure can affect the overall ENC and that this index may not be reliable for illustrating the respective contributions of selection and mutation to structuring codon usage patterns (Chamberlain et al., 2005).

Natural selection and mutation pressure are two of the main influencing factors that are thought to have a significant impact on codon usage patterns (Chakraborty et al., 2009). The majority of the codon usage among some RNA viruses would undoubtedly be explained by a general mutation pressure that affects the entire genome. Previous studies suggested that the base compositions at the third positions of the codon, mutational bias is mostly explained, whilst base compositions at the first and second positions, selective pressure were mostly validated (van Hemert and Berkhout, 2016). A neutrality plot analysis was carried out to determine the predominant factor, whether natural selection or mutation pressure influencing the CUB of LASV, using codon GC content as a basis and the results revealed that both mutational pressure and natural selection are responsible for shaping the CUB in LASV genomes. This phenomenon is similar to those given in the previous studies (Rahman et al., 2022c; Singh et al., 2016). Additionally, the PR2 plot analysis was performed to determine the factors influencing the codon usage bias in the LASV. According to Zhang et al. (2022), CUB may have resulted from mutational pressure alone if GC and AT were used proportionately; however, if GC and AT were not used proportionately, CUB may have been influenced by both natural selection and mutation. Here, the result suggested that, in addition to the mutation, natural selection may have influenced the patterns of codon usage in LASV, supporting the result of earlier studies (Rahman et al., 2022a; Uddin et al., 2018, 2020).

Codon adaptation index (CAI) is commonly utilized as an indicator to assess both codon usage bias and the extent to which viral genes have acclimated to their host organisms. Higher CAI values correspond to a greater prevalence of the most common codons and higher levels of gene expression within the host organism (Paola et al., 2018). Here, the high CAI value tendency of the hosts (*Homo sapiens* and *Mastomyous Natalensis)* suggests that selection pressure from *hosts* can impact the LASV codon usage, and that the codon usage evolution in LASV permits it to use the translation machinery of *hosts (Homo sapiens* and *Mastomyous Natalensis)* more capably. These results are in line with prior research, underscoring the importance of natural selection in impacting the CUB of viral genes (Khandia et al., 2019). CO analysis describes the discrepancy in the usage of codons. The primary aim of this analysis was to detect patterns in the variation of codon usage, and it was conducted by employing the RSCU values of 59 sense codons (Rahman et al., 2023; Zhang et al., 2011). In the plot of CO analysis, axis 1 and 2 are the two main factors of general discrepancy (Sharp and Li, 1986). Our results propose that the first axis signifies the virus strains, and the second axis signifies the countries where the virus arises. Using CO analysis, the LASV isolates were found in scattered form (Fig. 3), which shows that geographic regions are important for the evolution of LASV and the pattern of synonymous codons, and they may be used in the future to identify the origin of new LASV strains. Additionally, recent findings show that each affected country has more than one genetic lineage that is most widespread (Mercatelli and Giorgi, 2020). A phylogenetic study was used to determine how evolutionary processes have affected the codon usage pattern of LASV, which shows similarity to the pattern observed in COA van (Hemert, F and Berkhout, B. 2016). This result suggested that Lassa virus might be mutated due to some specific geographical effects (climatic changes, environmental changes etc.) (Chamberlain et al., 2005; Hewson et al., 2004).

5. Conclusion

The analysis of codon usage provides valuable information about the molecular evolution of viruses and their individual genes. This study has shown that both mutation pressure and natural selection have had a substantial influence on shaping the codon usage patterns of these strains. This study represents the first instance of codon usage analysis conducted on LASV strains. Further studies will be required to establish viral adaptation in various aspects and hosts.

CRediT authorship contribution statement

Siddiq Ur Rahman: Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Yikui Hu:** Writing – original draft, Writing – review & editing. **Hassan Ur Rehman:** Writing – original draft, Validation, Methodology, Formal analysis, Data curation. **May M. Alrashed:** Writing – review & editing. **Kotb A. Attia:** Writing – review & editing. **Ubaid Ullah:** Methodology, Formal analysis, Data curation. **Huiying Liang:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of competing interest

All the authors declare that they have no conflict of interest.

Acknowledgments and Funding

The authors extend their appreciation to the Researchers Supporting Project number (RSP-2025R-369), King Saud University, Riyadh, Saudi Arabia.

The authors would like to extend their sincere appreciation to the Research on the Identification Theory and Algorithm of Unknown Microorganisms, QN2023030008L, Ministry of Science and Technology, Foreign Youth Talent Program.

The second author of the manuscript is immensely grateful to Natural Science Foundation of Hubei Province of China (2017CFB565) and Research Project of Wuhan Municipal Health Commission (WX20C40) for their useful input.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2025.199528.

Data availability

Data will be made available on request.

References

Behura, S.K., Severson, D.W., 2013. Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. Biolog. Rev. 88 (1), 49–61. https://doi.org/10.1111/j.1469-185X.2012.00242.x.

Bera, B.C., Virmani, N., Kumar, N., et al., 2017. Genetic and codon usage bias analyses of polymerase genes of equine influenza virus and its relation to evolution. BMC Genom. 18 (1), 652. https://doi.org/10.1186/s12864-017-4063-1.

Bowen, G.S., Tomori, O., Wulff, H., et al., 1975. Lassa fever in Onitsha, East Central State, Nigeria in 1974. Bull. World Health Organ. 52 (4–6), 599–604.

Buchmeier, M.J., Southern, P.J., Parekh, B.S., et al., 1987. Site-specific antibodies define a cleavage site conserved among arenavirus GP-C glycoproteins. J. Virol. 61 (4), 982–985. https://doi.org/10.1128/JVI.61.4.982-985.1987.

Buckley, S.M., Casals, J., Downs, W.G., 1970. Isolation and antigenic characterization of lassa virus. Nature 227 (5254), 174. https://doi.org/10.1038/227174a0.

Burns, C.C., Shaw, J., Campagnoli, R., et al., 2006. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the Capsid Region. J. Virol. 80 (7), 3259–3272. https://doi.org/10.1128/JVI.80.7.3259.

Butt, A.M., Nasrullah, I., Tong, Y., 2014. Genome-wide analysis of codon usage and influencing factors in Chikungunya viruses. PLoS One 9 (3). https://doi.org/10.1371/journal.pone.0090905.

Butt, A.M., Nasrullah, I., Qamar, R., et al., 2016. Evolution of codon usage in Zika virus genomes is host and vector speci Fi c. Nat. Publish. Gr. 5 (August), e107. https://doi.org/10.1038/emi.2016.106.

Carey, D.E., Kemp, G.E., White, H.A., et al., 1972. Lassa fever. epidemiological aspects of the 1970 epidemic, Jos, Nigeria. Trans. R Soc. Trop. Med. Hyg. 66 (3), 402–408. https://doi.org/10.1016/0035-9203(72)90271-4.

Chakraborty, A., Uechi, T., Higa, S., Torihara, H., Kenmochi, N., 2009. Loss of ribosomal protein L11 affects zebrafish embryonic development through a P53-Dependent Apoptotic Response. PLoS ONE 4, e4152.

Chamberlain, J., Cook, N., Lloyd, G., et al., 2005. Co-evolutionary patterns of variation in small and large RNA segments of crimean-congo hemorrhagic fever virus. J. Gener. Virol. 86 (12), 3337–3341. https://doi.org/10.1099/vir.0.81213-0.

Costafreda, M.I., Pérez-Rodriguez, F.J., D'Andrea, L., et al., 2014. Hepatitis A virus adaptation to cellular shutoff is driven by dynamic adjustments of codon usage and results in the selection of populations with altered capsids. J. Virol. 88 (9), 5029–5041. https://doi.org/10.1128/JVI.00087-14.

Cristina, J., Fajardo, A., Sonora, M., et al., 2016. A detailed comparative analysis of codon usage bias in Zika virus. Virus Res. 223, 147–152. https://doi.org/10.1016/j.virusres.2016.06.022.

Datta, S.D., Talwar, A., Lee, J.T., 2020. A Proposed Framework and Timeline of the Spectrum of Disease Due to SARS-CoV-2 Infection: illness Beyond Acute Infection and Public Health Implications. JAMA 324 (22), 2251–2252. https://doi.org/10.1001/jama.2020.22717.

Deb, B., Uddin, A., Chakraborty, S., 2020. Codon Usage Pattern and Its Influencing Factors in Different Genomes of Hepadnaviruses. Arch. Virol. 165 (3), 557–570. https://doi.org/10.1007/s00705-020-04533-6.

Deb, B., Uddin, A., Chakraborty, S., 2021. Composition, Codon Usage Pattern, Protein Properties, and Influencing Factors in the Genomes of Members of the Family Anelloviridae. Arch. Virol. 166 (2), 461–474. https://doi.org/10.1007/s00705-020-04890-2.

Dixon, L., Tracy, S.K., Guilliland, K., et al., 2013. Outcomes of physiological and active third stage labour care amongst women in New Zealand. Midwifery 29 (1), 67–74. https://doi.org/10.1016/j.midw.2011.11.003.

Frame, J.D., Baldwin, J.M.J., Gocke, D.J., et al., 1970. Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. Am. J. Trop. Med. Hyg. 19 (4), 670–676. https://doi.org/10.4269/ajtmh.1970.19.670.

Frame, J.D., 1975. Surveillance of Lassa Fever in Missionaries Stationed in West Africa. Bull. World Health Organ. 52 (4–6), 593–598.

Fraser, D.W., Campbell, C.C., Monath, T.P., et al., 1974. Lassa fever in the eastern province of Sierra Leone, 1970-1972. I. Epidemiologic studies. Am. J. Trop. Med. Hyg. 23 (6), 1131–1139. https://doi.org/10.4269/ajtmh.1974.23.1131.

Goeijenbier, M., Wagenaar, J., Goris, M., et al., 2013. Rodent-borne hemorrhagic fevers: under-recognized, widely spread and preventable – epidemiology, diagnostics and treatment. Crit. Rev. Microbiol. 39 (1), 26–42. https://doi.org/10.3109/1040841X.2012.686481.

Hewson, R., Gmyl, A., Gmyl, L., et al., 2004. Evidence of segment reassortment in Crimean-Congo haemorrhagic fever virus. J. Gener. Virol. 85 (10), 3059–3070. https://doi.org/10.1099/vir.0.80121-0.

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), 1–7. https://doi.org/10.1016/S0168-1702(02)00309-X.

Johnson, K.M., Taylor, P., Elliott, L.H., et al., 1981. Recovery of a Lassa-Related Arenavirus in Zimbabwe. Am. J. Trop. Med. Hyg. 30 (6), 1291–1293. https://doi.org/10.4269/ajtmh.1981.30.1291.

Karniychuk, U.U., 2016. Analysis of the Synonymous Codon Usage Bias in Recently Emerged Enterovirus D68 Strains. Virus Res. 223, 73–79.

Kattoor, J.J., Malik, Y.S., Sasidharan, A., et al., 2015. Analysis of codon usage pattern evolution in Avian rotaviruses and their preferred host. Infect. Genet. Evol. 34, 17–25. https://doi.org/10.1016/j.meegid.2015.06.018.

Keenlyside, R.A., McCormick, J.B., Webb, P.A., et al., 1983. Case-control study of Mastomys Natalensis and humans in lassa virus-infected households in Sierra Leone. Am. J. Trop. Med. Hyg. 32 (4), 829–837. https://doi.org/10.4269/ajtmh.1983.32.829.

Khandia, R., Singhal, S., Kumar, U., et al., 2019. Analysis of Nipah Virus Codon Usage and Adaptation to Hosts. Front. Microbiol. 10, 886. https://doi.org/10.3389/fmicb.2019.00886.

Kumar, N., Bera, B.C., Greenbaum, B.D., et al., 2016a. Revelation of influencing factors in overall codon usage bias of equine influenza viruses. PLoS One 11 (4), e0154376. https://doi.org/10.1371/journal.pone.0154376.

Kumar, S., Stecher, G., Tamura, K., 2016b. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874. https://doi.org/10.1093/molbev/msw054.

Li, G., Chen, X., Li, X., et al., 2023. Analyzing the evolution and host adaptation of the rabies virus from the perspective of codon usage bias. Chaintoutis SC. ed Transbound. Emerg. Dis. 2023, 4667253. https://doi.org/10.1155/2023/4667253.
Liu, H., Rahman, S.U., Mao, Y., et al., 2017. Codon usage bias in 5' terminal coding sequences reveals distinct enrichment of gene functions. Genomics (July). https://doi.org/10.1016/j.ygeno.2017.07.008, 0–1.
Lukashevich, I.S., Stelmakh, T.A., Golubev, V.P., et al., 1984. Ribonucleic acids of machupo and Lassa viruses. Arch. Virol. 79 (3–4), 189–203. https://doi.org/10.1007/BF01310811.
Ma, M.-R., Hui, L., Wang, M.-L., et al., 2015. Synonymous codon selection in the hepatitis B virus translation initiation region. Genet. Mol. Res. 14 (3), 8955–8963. https://doi.org/10.4238/2015.August.7.4.
McCormick, J.B., Fisher-Hoch, S.P., 2002. Lassa fever. Curr. Top. Microbiol. Immunol. 262, 75–109. https://doi.org/10.1007/978-3-642-56029-3_4.
McCormick, J.B., Webb, P.A., Krebs, J.W., et al., 1987. A prospective study of the epidemiology and ecology of Lassa fever. J. Infect. Dis. 155 (3), 437–444.
Mercatelli, D., Giorgi, F.M., 2020. Geographic and genomic distribution of SARS-CoV-2 mutations. Front. Microbiol. 11 (July), 1–13. https://doi.org/10.3389/fmicb.2020.01800.
Monath, T.P., Mertens, P.E., Patton, R., et al., 1973. A hospital epidemic of Lassa fever in Zorzor, Liberia, March-April 1972. Am. J. Trop. Med. Hyg. 22 (6), 773–779. https://doi.org/10.4269/ajtmh.1973.22.773.
Monath, T.P., Newhouse, V.F., Kemp, G.E., et al., 1974. Lassa virus isolation from Mastomys Natalensis rodents during an epidemic in Sierra Leone. Science 185 (4147), 263–265. https://doi.org/10.1126/science.185.4147.263.
Moratorio, G., Iriarte, A., Moreno, P., et al., 2013. A detailed comparative analysis on the overall codon usage patterns in West Nile Virus. Infect. Genet. Evolut. 14 (1), 396–400. https://doi.org/10.1016/j.meegid.2013.01.001.
Mueller, S., Papamichail, D., Coleman, J.R., et al., 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 (19), 9687–9696. https://doi.org/10.1128/JVI.00738-06.
Nasrullah, I., Butt, A.M., Tahir, S., et al., 2015. Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on marburg virus evolution. BMC Evol. Biol. 15 (1), 174. https://doi.org/10.1186/s12862-015-0456-4.
Nguyen, T.H., Wang, D., Rahman, S.U., et al., 2021. Analysis of codon usage patterns and influencing factors in rice Tungro Bacilliform virus. Infect. Genet. Evolut. 90 (November 2020), 104750. https://doi.org/10.1016/j.meegid.2021.104750.
Paola, N.D., Freire, C.C.d.M., Zanotto, P.M.d.A., 2018. Does adaptation to vertebrate codon usage relate to flavivirus emergence potential? PLOS ONE 13 (1), e0191652. https://doi.org/10.1371/journal.pone.0191652.
Puigbò, P., Bravo, I.G., Garcia-Vallve, S., 2008. CAIcal: a combined set of tools to assess codon usage adaptation. Biol. Direct. 3, 38. https://doi.org/10.1186/1745-6150-3-38.
Quax, T.E.F., Claassens, N.J., Söll, D., et al., 2015. Codon bias as a means to fine-tune gene expression. Mol. Cell. 59 (2), 149–161. https://doi.org/10.1016/j.molcel.2015.05.035.
Rahman, S.U., Yao, X., Li, X., et al., 2018. Analysis of codon usage bias of Crimean-Congo hemorrhagic fever virus and its adaptation to hosts. Infect. Genet. Evolut. 58 (November 2017), 1–16. https://doi.org/10.1016/j.meegid.2017.11.027.
Rahman, S.U., Abdullah, M., Khan, A.W., et al., 2022a. A detailed comparative analysis of codon usage bias in Alongshan virus. Virus Res. 308 (November 2021), 198646. https://doi.org/10.1016/j.virusres.2021.198646.
Rahman, S.U., Nawaz, S., Ullah, S., et al., 2022b. Study of codon usage patterns and influencing factors in rice yellow mottle virus based on coding sequence data. Agronomy 12 (9), 1990. https://doi.org/10.3390/agronomy12091990.
Rahman, S.U., Rehman, H.U., Rahman, I.U., et al., 2022c. Analysis of codon usage bias of lumpy skin disease virus causing livestock infection. Front. Vet. Sci. 9. https://doi.org/10.3389/fvets.2022.1071097.
Rahman, S.U., Rehman, H.U., Rahman, I.U., et al., 2023. Evolution of codon usage in Taenia Saginata genomes and its impact on the host. Front. Vet. Sci. 9. https://doi.org/10.3389/fvets.2022.1021440.
Sharma, O., Sultan, A.A., Ding, H., et al., 2020. A review of the progress and challenges of developing a vaccine for COVID-19. Front. Immunol. 11, 585354. https://doi.org/10.3389/fimmu.2020.585354.
Sharp, P.M., Li, W.H., 1986. Codon usage in regulatory genes in Escherichia Coli does not reflect selection for "rare" codons. Nucl. Acid. Res. 14 (19), 7737–7749. https://doi.org/10.1093/nar/14.19.7737.
Singh, N.K., Tyagi, A., Kaur, R., et al., 2016. Characterization of codon usage pattern and influencing factors in Japanese encephalitis virus. Virus Res. 221, 58–65. https://doi.org/10.1016/j.virusres.2016.05.008.
Sueoka, N., 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. U.S.A. 85 (8), 2653–2657. https://doi.org/10.1073/pnas.85.8.2653.
Sueoka, N., 1992. Directional mutation pressure, selective constraints, and genetic equilibria. J. Mol. Evol. 34, 95–114. https://doi.org/10.1007/BF00182387.
Sun, J., Zhao, W., Wang, R., et al., 2020. Analysis of the codon usage pattern of HA and NA genes of H7N9 Influenza A virus. Int. J. Mol. Sci. 21 (19). https://doi.org/10.3390/ijms21197129.
Tao, P., Dai, L., Luo, M., et al., 2009. Analysis of synonymous codon usage in classical Swine fever virus. Virus Gene. 38 (1), 104–112. https://doi.org/10.1007/s11262-008-0296-z.
Tyagi, N., Sardar, R., Gupta, D., 2022. Natural selection plays a significant role in governing the codon usage bias in the novel SARS-CoV-2 variants of concern (VOC). PeerJ 10, e13562. https://doi.org/10.7717/peerj.13562.
Uddin, A., Mazumder, T.H., Chakraborty, S., 2018. Understanding molecular biology ofcodon usage in mitochondrial complex IV genes of electron transport system: relevance to mitochondrial diseases. J. Cell. Physiol. 1–17. https://doi.org/10.1002/jcp.27375.
Uddin, A., 2020. Compositional features and codon usage pattern of genes associatedwith anxiety in human. Mol. Neurobiol. https://doi.org/10.1007/s12035-020-02068-0.
van Hemert, F., Berkhout, B., 2016. Nucleotide composition of the Zika virus RNA genome and its codon usage. Virol. J. 13 (1), 95. https://doi.org/10.1186/s12985-016-0551-1.
Vieth, S., Torda, A.E., Asper, M., et al., 2004. Sequence analysis of L RNA of Lassa virus. Virology 318 (1), 153–168. https://doi.org/10.1016/j.virol.2003.09.009.
Wang, F., Zhang, N., Zhao, C., et al., 2023a. Codon usage bias analysis of mitochondrial protein-coding genes in 12 species of candida. J. Genet. 102.
Wang, Y., Jiang, D., Guo, K., et al., 2023b. Comparative analysis of codon usage patterns in chloroplast genomes of ten epimedium species. BMC Genom. Data 24 (1), 3. https://doi.org/10.1186/s12863-023-01104-x.
Wong, E.H.M., Smith, D.K., Rabadan, R., et al., 2010. Codon usage bias and the evolution of Influenza A viruses. codon usage biases of Influenza virus. BMC Evol. Biol. 10, 253. https://doi.org/10.1186/1471-2148-10-253.
Wright, F., 1990. The "effective number of codons" used in a gene. Gene 87 (1), 23–29. https://doi.org/10.1016/0378-1119(90)90491-9.
Wulff, H., McIntosh, B.M., Hamner, D.B., et al., 1977. Isolation of an arenavirus closely related to lassa virus from mastomys natalensis in South-East Africa. Bull. World Health Organ. 55 (4), 441–444.
Yang, C., Zhao, Q., Wang, Y., et al., 2021. Comparative analysis of genomic and transcriptome sequences reveals divergent patterns of codon bias in wheat and its ancestor species. Front. Genet. 12, 732432. https://doi.org/10.3389/fgene.2021.732432.
Yao, X., Fan, Q., Yao, B., et al., 2020. Codon usage bias analysis of bluetongue virus causing livestock infection. Front. Microbiol. 11 (May), 1–12. https://doi.org/10.3389/fmicb.2020.00655.
Zhang, J., Wang, M., Liu, W., et al., 2011. Analysis of codon usage and nucleotide composition bias in polioviruses. Virol. J. 8 (1), 146. https://doi.org/10.1186/1743-422X-8-146.
Zhang, Y., Shen, Z., Meng, X., et al., 2022. Codon usage patterns across seven rosales species. BMC Plant Biol. 22, 65.