



Characteristics of synonymous codon usage bias in the beginning region of West Nile virus

X.X. Ma*, Y.P. Feng*, J.L. Liu, L. Chen, Y.Q. Zhao, P.H. Guo, J.Z. Guo, L.Y. Ma, and Z.R. Ma

Key Laboratory of Bioengineering & Biotechnology of State Ethnic Affairs Commission, Engineering & Technology Research Center for Animal Cell, College of Life Science and Engineering, Northwest University for Nationalities, Lanzhou, China

*These authors contributed equally to this study.

Corresponding author: Z.R. Ma

E-mail: maxiaoxia956@163.com / mZR@xbmu.edu.cn

Genet. Mol. Res. 13 (3): 7347-7355 (2014)

Received July 8, 2013

Accepted December 3, 2013

Published September 5, 2014

DOI <http://dx.doi.org/10.4238/2014.September.5.13>

ABSTRACT. Adaptation in the overall codon usage pattern of West Nile virus (WNV) to that of two hosts was estimated based on the synonymous codon usage value (RSCU). Synonymous codon usage biases for the beginning coding sequence of this virus were also analyzed by calculating the usage fluctuation for each synonymous codon along the target region (the first 270 codon sites of the whole coding sequence of WNV). Adaptation of WNV to *Anopheles gambiae* regarding the overall codon usage revealed a mixture of synonymous codon usage patterns between this virus and its vector. Regarding the adaptation of WNV to its dead-end host and codon usage, although a mixture of overall codon usage patterns exists, the number of codons with reversed tendency codon usage is lower than that between the virus and its vector. In addition, some codons with low RSCU values for this virus are highly selected in the beginning translation region of WNV, while codons with low RSCU values in this region tend to pair with tRNAs present in low abundance in the host, suggesting that highly selected

codons in a specific region in the beginning region of WNV are, to some degree, influenced by the corresponding low tRNA abundance of hosts to regulate the translation speed of the WNV polyprotein.

Key words: West Nile virus; Codon usage; tRNA abundance; Beginning translation region

INTRODUCTION

West Nile virus (WNV) was first isolated in a woman in the West Nile district of Uganda in 1937. Currently, this virus exists in Africa, Europe, the Middle East, North America, and West Asia (Maeda et al., 2009). WNV belongs to a member of the flavivirus genus and has a positive and single-stranded RNA genome, from which ten mature viral proteins are produced via proteolytic processing of the single polyprotein by the viral serine protease and various cellular proteases (Nowak et al., 1989; Brinton, 2002). In general, WNV is maintained in nature in a cycle involving transmission between birds and mosquitoes, but humans and other mammals can also be infected through bites from infected mosquitoes. Humans became incidentally infected and develop neurological disease or death depending on the virus type, virulence, age, immune status, and genetic factors of the infected individual (Mackenzie et al., 1994; Wu et al., 2000; Beasley et al., 2004). Regarding WNV tropism, the virus is neurotropic *in vivo* and infects some visceral organs (Diamond et al., 2003). Although WNV is a significant global human health problem, studies examining the molecular interaction of this virus with its natural host are limited (Brinton, 2002). Krishnan et al. (2008) used a human-genome-wide RNAi screen to identify 305 host proteins potentially involved in WNV infection and found a complex dependence of WNV on the host cell physiology, which required a wide variety of molecules and cellular pathways for successful infection. In addition to analyzing pathways of virus infection in hosts, examining the overall codon usage for an RNA virus open reading frame (ORF) is useful for estimating the evolutionary processes and genetic features of viruses that respond or adapt to the host cell environment (Bahir et al., 2009; Lobo et al., 2009; Wong et al., 2010; Zhou et al., 2013a,b). Recent progress has been made in measuring viral protein translation at the level of synonymous codon usage bias (Sanchez et al., 2003; Burns et al., 2006; Mueller et al., 2006; Bahir et al., 2009; Wong et al., 2010; Poyry and Jackson, 2011; Zhou et al., 2011); however, the contribution of the WNV coding sequence to replication efficiency remains unclear. Here, we employed a simple method for analyzing synonymous codon usage bias in the beginning region of the WNV coding sequence to evaluate the role of codon usage in regulating WNV polyprotein translation.

MATERIAL AND METHODS

Sequence data of WNV and codon usage frequency data of hosts

A total of 130 coding sequences of WNV were downloaded from the National Center for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov/Genbank/>) (listed in [Table S1](#)). Synonymous codon usage data of humans and *Anopheles gambiae* were obtained from the codon usage database at <http://www.kazusa.or.jp/codon/> (Nakamura et al., 2000). In addition, to evaluate the relationship between synonymous codon usage bias and

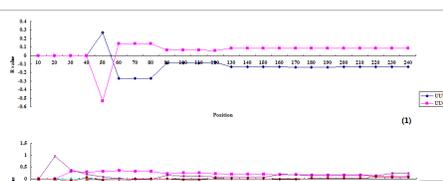
corresponding tRNA concentrations, tRNA anticodon information for humans and *Anopheles gambiae* was obtained from the Genomic tRNA Database (<http://gtrnadb.ucsc.edu/>) and is listed in [Table S2](#).

Comparison of overall codon usage between WNV and hosts

Since relative synonymous codon usage (RSCU) values are not affected by amino acid composition and coding sequence size, we employed the method reported by Sharp and Li (1986) to calculate RSCU data of WNV and its two hosts and to evaluate viral adaptation to its hosts at the codon usage level.

Calculation of synonymous codon usage bias in the beginning coding sequence of WNV

To analyze synonymous codon usage bias in the beginning coding sequence (720 nt) of WNV, we developed a simple method based on a previous report (Zhou et al., 2011), which involved the following equation:



where n_x is the sum of a synonymous codon in the target beginning coding sequence of WNV, N_x is the sum of the corresponding amino acids in the corresponding region, 'x' represents different sizes of the beginning region ranging from the first codon to the 240th codon, which can be 10, 20, 30, 40... 210, 220, 230, or 240, n is the sum of this synonymous codon in the entire coding sequence of WNV, and N is the sum of the corresponding amino acid in the whole coding sequence. Note that if one codon exists in the specific target region in this study, the R value for this codon is defined as zero in the specific region.

RESULTS

Comparative analysis of overall codon usage pattern between WNV and its hosts

A mixture of synonymous codons is used between this virus and its vector in the adaptation of WNV to *Anopheles gambiae* at the level of overall codon usage. Similar-usage codons include all synonymous codons for Phe, Ile, Val, Tyr, His, Gln, Asn, Lys, Asp, Glu, Cys, and UUA, CUU, CUC, and CUA for Leu, AGC for Ser, CCC for Pro, ACC for Thr, GCC and GCA for Ala, CGA for Arg, and GGG for Gly; the remaining codons have a reversed tendency for codon usage (Table 1). Although a mixture of overall codon usage pattern exists in the adaptation of WNV to its dead-end host at the codon usage level, the number of codons with a reversed tendency for codon usage is lower than that between the virus and its vector. The reversed codon usage pattern includes UUG and CUG for Leu, UCU, UCC, UCA, UCG, and AGU for Ser, CCC and CCA for Pro, GCC for Ala, CGG and AGA for Arg, and GGA for Gly (Table 1).

Table 1. Comparison of the synonymous codon usage pattern between West Nile virus (WNV) and its hosts.

	WNV	<i>Anopheles gambiae</i>	Human
UUU(F)	0.90	0.69	0.87
UUC(F)	1.10	1.31	1.13
UUA(L)	0.19	0.28	0.39
UUG(L)	1.46	0.73	0.73
CUU(L)	0.72	0.52	0.73
CUC(L)	1.18	0.94	1.22
CUA(L)	0.66	0.46	0.40
CUG(L)	1.78	3.06	2.53
AUU(I)	1.08	0.79	1.04
AUC(I)	1.30	1.72	1.52
AUA(I)	0.62	0.49	0.44
GUU(V)	0.77	0.65	0.69
GUC(V)	1.07	0.94	1.00
GUA(V)	0.32	0.51	0.42
GUG(V)	1.83	1.90	1.90
UCU(S)	0.69	0.38	1.11
UCC(S)	0.68	1.09	1.39
UCA(S)	1.57	0.48	0.84
UCG(S)	0.57	1.79	0.33
AGU(S)	1.12	0.65	0.84
AGC(S)	1.38	1.61	1.50
CCU(P)	0.87	0.36	1.12
CCC(P)	0.90	0.83	1.35
CCA(P)	1.79	0.75	1.07
CCG(P)	0.45	2.06	0.46
ACU(T)	0.84	0.42	0.94
ACC(T)	1.27	1.31	1.52
ACA(T)	1.23	0.55	1.07
ACG(T)	0.65	1.71	0.46
GCU(A)	1.32	0.62	1.09
GCC(A)	1.17	1.38	1.64
GCA(A)	0.99	0.71	0.85
GCG(A)	0.52	1.29	0.42
UAU(Y)	0.75	0.50	0.84
UAC(Y)	1.25	1.50	1.16
CAU(H)	0.80	0.73	0.81
CAC(H)	1.20	1.27	1.19
CAA(Q)	0.88	0.52	0.51
CAG(Q)	1.12	1.48	1.49
AAU(N)	0.74	0.64	0.89
AAC(N)	1.26	1.36	1.11
AAA(K)	0.84	0.60	0.82
AAG(K)	1.16	1.40	1.18
GAU(D)	0.94	0.94	0.89
GAC(D)	1.06	1.06	1.11
GAA(E)	0.98	0.76	0.81
GAG(E)	1.02	1.24	1.19
UGU(C)	0.90	0.78	0.86
UGC(C)	1.10	1.22	1.14
CGU(R)	0.55	0.95	0.51
CGC(R)	0.85	1.93	1.20
CGA(R)	0.52	0.78	0.63
CGG(R)	0.73	1.56	1.20
AGA(R)	2.00	0.42	1.20
AGG(R)	1.34	0.36	1.26
GGU(G)	0.44	1.05	0.64
GGC(G)	0.81	1.52	1.40
GGA(G)	2.01	0.80	0.98
GGG(G)	0.74	0.63	0.98

Interestingly, 45 of 59 synonymous codons have a generally similar codon usage pattern among the virus, its vector, and its dead-end host. These results suggest that the overall evolutionary

characteristics of the codon usage pattern for WNV are influenced by its vectors and dead-end hosts. However, some amino acids (Ser, Pro, Thr, Arg, and Gly), have more than 2 synonymous codons and show reversed codon usage to a large degree compared with the corresponding amino acids for the vector and the dead-end host. This indicates that WNV presents its specific genetic features depending on synonymous codon usage patterns to perform its biological function in different cell types. Notably, CUG for Leu, CCA for Pro, and AGA and GGA for Arg are highly selected by WNV and the three codons (excluding CUG; this codon usage is influenced by the hosts) are selected at relatively low levels by the two organisms and correspond to relatively low tRNA densities. These results suggest that WNV largely uses these codons to pair with tRNA present in low abundance in order to regulate its polyprotein translation.

Synonymous codon usage bias in the beginning coding sequence of WNV

Regarding the tendency of synonymous codon usage in the beginning coding sequence of WNV, the 59 synonymous codons are not selected at an equal frequency in target regions of different scales ([Figure S1](#)). Although WNV is an RNA virus and has a high mutation rate, some codons show a bias for codon usage; UAU and UAC for Tyr, CAU and CAC for His, CAG for Gln, GAA and GAG for Glu, and UGU and UGC for Cys are generally not present in the beginning region of the virus. In addition, some codons show an obvious reversed bias against the overall codon usage pattern of WNV. For the Phe codons, although the RSCU value for UUU is lower than that for UUC (Table 1), UUU appears at a higher frequency than UUC in the first 50 codons ([Figure S1-1](#)). For Leu, selection of CUA in the first 20 codons is stronger than at other synonymous codons ([Figure S1-2](#)), but the RSCU value for CUA is relatively low in the synonymous family (Table 1). For Ile, although the RSCU value for AUC is the highest in the synonymous family (Table 1), this codon appears at the lowest frequency in the first 20 codons ([Figure S1-3](#)). For Val, although the RSCU value for GUG is the highest in the codon family (Table 1), this codon does not exist in the first 40 codons of the beginning region ([Figure S1-4](#)). For Ser, UCU with a low RSCU value occurs at a high frequency in the first 10 codons in the beginning region, AGU with a high RSCU value appears infrequently across the first 80 codons, and UCA with the highest RSCU value in the codon family appears very infrequently in the first 90 codons (Table 1 and [Figure S1-5](#)). For Pro, CCA, with the highest RSCU value in the codon family, is weakly selected across the first 170 codons in the beginning region, while CCG, with the lowest RSCU value, is strongly selected, particularly from the 70th codon to the 170th codon (Table 1 and [Figure S1-6](#)). For Thr, ACC and ACA, with high RSCU values, are not found in the first 40 codons, but ACG, with the lowest RSCU value, is the most frequently selected in the first 40 codons (Table 1 and [Figure S1-7](#)). For Ala, GCC, with a high RSCU value, was selected with the lowest frequency in the beginning region (Table 1 and [Figure S1-8](#)). For Tyr, the two codons are never found in the first 160 codons ([Figure S1-9](#)). For His, the two codons are never found in the first 70 codons, and their usage pattern in the beginning region is similar to the overall codon usage pattern (Table 1 and [Figure S1-10](#)). For Gln, although the RSCU value for CAA is lower than that of CAG (Table 1), the former has a stronger tendency than the latter to be found in the beginning region ([Figure S1-11](#)). For Asn, the RSCU value of AAU is lower than that of AAC (Table 1), but the former is more frequently selected in the beginning region than the latter ([Figure S1-12](#)). For Lys, the two codons have a similar codon usage pattern in the beginning region ([Figure S1-13](#)). For Asp, GAU has a low RSCU value and is more frequently

selected than GAC in the first 30 codons (Table 1 and [Figure S1-14](#)). For Glu, although the two codons have similar RSCU values and are not found in the first 60 codons, GAA is selected at a relatively high frequency in the beginning region (Table 1 and [Figure S1-15](#)). For Cys, the two codons do not appear in the first 110 codons, but UGU, with a relatively low RSCU value, is selected more frequently than UGC in the range of the 120th codon to the 150th codon, after which the situation reverses from the 160th codon to the 240th codon (Table 1 and [Figure S1-16](#)). For Arg, AGA, with the highest RSCU value in the codon family, is not preferentially selected in the beginning region, while CGC, CGG, and CGA, with low RSCU values, show a clear tendency to be found in the beginning region compared with AGA and AGG (Table 1 and [Figure S1-17](#)). For Gly, GGA with the highest RSCU value in the codon family is not the preferred codon for the beginning region, while GGG is generally selected at a high frequency in this region (Table 1 and [Figure S1-18](#)). These results suggest that several codons with low RSCU values are frequently selected in the beginning region compared with those with high RSCU values; these non-preferred codons may play a role in regulating the translation of WNV polyproteins by reducing the translation speed.

Association between codon usage bias in the beginning region and tRNA pool in the host

To evaluate the potential influence of tRNA concentration on ribosome movement, we analyzed the specific tRNA concentration for the preferred corresponding codon chosen at the beginning region of WNV. Based on identified tRNA abundance ([Table S2](#)), we found that some non-preferred codons with a strong tendency to be found in the beginning region of WNV fail to correspond to a rare tRNA concentration. For Leu, CUA, corresponding to low tRNA density in its vector and dead-end host, is highly selected in the first 20 codons, while CUG, corresponding to high tRNA density, showed no obvious tendency to be found at high or low levels in the beginning region ([Table S2](#) and [Figure S1-2](#)). For Ile, although AUU, corresponding to high tRNA density, showed a clear tendency to be found in the first 30 codons, the selection of AUA pairing to relatively low tRNA density is biased toward a specific region from the 50th codon to the 90th codon ([Table S2](#) and [Figure S1-2](#)). For Val, although there is no information about tRNA density pairing to GUC, GUG, corresponding to high tRNA density, was not observed in the first 40 codons, and GUU, with high tRNA density, is weakly selected in the beginning region ([Table S2](#) and [Figure S1-4](#)).

For Ser, different tRNA abundance levels pairing to synonymous codons may play a slight role in translation speed, since UCA, with low tRNA density, is weakly selected and UCU and AGC, with high tRNA densities, are highly selected in the beginning region ([Table S2](#) and [Figure S1-5](#)). For Pro, because no information is available regarding tRNA abundance pairing to CCC, there is no clear trend for codons with low tRNA density to present in the beginning region ([Table S2](#) and [Figure S1-6](#)). For Thr, ACU, with high tRNA density, is weakly selected in the beginning region, and ACG and ACA, corresponding to relatively low tRNA densities, are strongly selected in the specific regions of the beginning region of WNV ([Table S2](#) and [Figure S1-7](#)). For Ala, GCU, with high tRNA density, is highly selected in the beginning region, but GCA and GCG, with relatively low tRNA densities, are rarely found in the first 70 codons ([Table S2](#) and [Figure S1-8](#)). For Tyr, based on the tRNA abundance for humans, the two codons are never present in the first 160 codons, and UAC, with high tRNA density, is

highly selected in the first 170 codons ([Table S2](#) and [Figure S1-9](#)). For Gln, CAA, with low tRNA density, is highly selected in the beginning region, but CAG, with high tRNA density, is weakly selected in this region ([Table S2](#) and [Figure S1-11](#)). For Asn, AAU, with low tRNA density, is selected at a higher level than AAC, with high tRNA density, in the beginning region ([Table S2](#) and [Figure S1-12](#)). For Lys, although the two codons have a similar usage pattern in the beginning region, AAA, with low tRNA density, is used in relatively high frequency in the first 30 codons ([Table S2](#) and [Figure S1-13](#)). For Glu, the two codons are never found in the first 60 codons, whereas GAA, with relatively low tRNA density, is more highly selected in the beginning region ([Table S2](#) and [Figure S1-15](#)). For Arg, CGU, with relatively high tRNA density, is never present in the beginning region, while AGG, with relatively low tRNA density, is generally found in the first 10 codons ([Table S2](#) and [Figure S1-17](#)). For Gly, based on tRNA densities pairing to GGG and GGC in humans, the former is more highly selected in the first 30 codons than the latter ([Table S2](#) and [Figure S1-18](#)). These results suggest that some codons that are highly selected in a specific region of the beginning region of WNV are influenced by corresponding low tRNA densities.

DISCUSSION

Two options for dealing with altered conditions are available to a virus, including: i) adaptation to available ecological niches and ii) extinction. Although WNV is a mosquito-borne flavivirus that primarily cycles in birds, it occasionally infects humans. Recently, the incidence of WNV outbreaks in humans has increased, and the infections have resulted in severe disease (Hayes and Gubler, 2006; Murray et al., 2010b; Barzon et al., 2012). The purpose of this study was to explore the potential effect of synonymous codon usage bias in the beginning region of WNV and the tRNA concentrations for this region in the vector and dead-end host. Our observations indicated that some patients were persistently infected with WNV and that the kidney may be a preferred site of continued replication and source of shedding (Murray et al., 2010a), suggesting that codon usage in this virus may be influenced by its dead-end host. The significant increase in translation rate owing to the appropriate arrangement of codons has an important impact on how the translation machinery functions. Much of this synonymous codon usage does match with the corresponding codons in the two organisms. This adaptation of the overall codon usage pattern of WNV to the cellular codon usage pattern may enable the virus to minimize the cost of protein expression. Optimization of the codon usage pattern is frequently required for efficient expression of genes in heterologous host systems (Kane, 1995; Smith, 1996; Andre et al., 1998; Yadava and Ockenhouse, 2003). Furthermore, Cannarozzi et al. (2010) found that the advantage to frequently used codons is the efficiency of reusing the corresponding tRNA. tRNA molecules exiting the ribosome remain associated with the translational machinery so that they are readily available when the next identical codon is detected (Fredrick and Ibba, 2010). Gene expression is governed at multiple stages, including translation, initiation, and elongation. The extent of adaptation of the coding sequence to cellular tRNA density is an important factor in the elongation phase of translation (dos Reis et al., 2004). Tuller et al. (2010) found that the first 30-50 codons are, on average, translated with low efficiency in order to regulate ribosome density along mRNA (Tuller et al., 2010). In this study, codons with low tRNA abundance showed a clear tendency to be found in this specific region and even beyond this region. This phenomenon also demonstrates that in order to regulate

translation of the viral polyprotein, the overall codon usage pattern of WNV can be influenced by the environment of codon usage of the host cell, since codon usage and tRNA regulation can co-evolve (Plotkin and Kudla, 2011). Slowing of translation elongation immediately after initiation would effectively generate more uniform spacing between ribosomes further down the mRNA, which should help to reduce the ribosome traffic jam and promote efficient protein synthesis for WNV. The ribosome traffic jam in gene expression reduces processivity and even results in ribosome stalling, and the ribosome density is inversely proportional to the elongation rate (Kim and Hollingsworth, 1992; Parmley and Huynen, 2009; Ehrenberg et al., 2010). Our observation that some non-preferred codons corresponding to low tRNA abundances have a strong tendency to be found in the beginning coding sequence of WNV implies that in order to efficiently use the translational resources of host cells, WNV may regularly adapt to its hosts with respect to a synonymous codon usage pattern in its beginning region.

ACKNOWLEDGMENTS

Research supported by grants from the National Natural Science Foundation of China (#31160033) and the central universities deriving from the Northwest University for Nationalities (#ZYZ2012067) and the university's scientific fund for the middle-aged and young scientists deriving from the Northwest University for Nationalities (#12XB37).

[Supplementary material](#)

REFERENCES

- Andre S, Seed B, Eberle J, Schraut W, et al. (1998). Increased immune response elicited by DNA vaccination with a synthetic gp120 sequence with optimized codon usage. *J. Virol.* 72: 1497-1503.
- Bahir I, Fromer M, Prat Y and Linial M (2009). Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Mol. Syst. Biol.* 5: 311.
- Barzon L, Pacenti M, Franchin E, Lavezzo E, et al. (2012). New endemic West Nile virus lineage 1a in northern Italy, July 2012. *Euro. Surveill.* 17:pii:20244.
- Beasley DW, Davis CT, Whiteman M, Granwehr B, et al. (2004). Molecular determinants of virulence of West Nile virus in North America. *Arch. Virol. Suppl.* 35-41.
- Brinton MA (2002). The molecular biology of West Nile Virus: a new invader of the western hemisphere. *Annu. Rev. Microbiol.* 56: 371-402.
- Burns CC, Shaw J, Campagnoli R, Jorba J, et al. (2006). Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. *J. Virol.* 80: 3259-3272.
- Cannarozzi G, Schraudolph NN, Faty M, von Rohr P, et al. (2010). A role for codon order in translation dynamics. *Cell* 141: 355-367.
- Diamond MS, Shrestha B, Marri A, Mahan D, et al. (2003). B cells and antibody play critical roles in the immediate defense of disseminated infection by West Nile encephalitis virus. *J. Virol.* 77: 2578-2586.
- dos Reis M, Savva R and Wernisch L (2004). Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32: 5036-5044.
- Ehrenberg M, Dennis PP and Bremer H (2010). Maximum rrn promoter activity in *Escherichia coli* at saturating concentrations of free RNA polymerase. *Biochimie* 92: 12-20.
- Fredrick K and Ibbas M (2010). How the sequence of a gene can tune its translation. *Cell* 141: 227-229.
- Hayes EB and Gubler DJ (2006). West Nile virus: epidemiology and clinical features of an emerging epidemic in the United States. *Annu. Rev. Med.* 57: 181-194.
- Kane JF (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* 6: 494-500.
- Kim JK and Hollingsworth MJ (1992). Localization of *in vivo* ribosome pause sites. *Anal. Biochem.* 206: 183-188.

- Krishnan MN, Ng A, Sukumaran B, Gilfoxy FD, et al. (2008). RNA interference screen for human genes associated with West Nile virus infection. *Nature* 455: 242-245.
- Lobo FP, Mota BE, Pena SD, Azevedo V, et al. (2009). Virus-host coevolution: common patterns of nucleotide motif usage in Flaviviridae and their hosts. *PLoS One*. 4: e6282.
- Mackenzie JS, Lindsay MD, Coelen RJ, Broom AK, et al. (1994). Arboviruses causing human disease in the Australasian zoogeographic region. *Arch. Virol.* 136: 447-467.
- Maeda A, Murata R, Akiyama M, Takashima I, et al. (2009). A PCR-based protocol for the generation of a recombinant West Nile virus. *Virus Res.* 144: 35-43.
- Mueller S, Papamichail D, Coleman JR, Skiema S, 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: 9687-9696.
- Murray K, Walker C, Herrington E, Lewis JA, et al. (2010a). Persistent infection with West Nile virus years after initial infection. *J. Infect. Dis.* 201: 2-4.
- Murray KO, Mertens E and Despres P (2010b). West Nile virus and its emergence in the United States of America. *Vet. Res.* 41: 67.
- Nakamura Y, Gojobori T and Ikemura T (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* 28: 292.
- Nowak T, Farber PM, Wengler G and Wengler G (1989). Analyses of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology* 169: 365-376.
- Parmley JL and Huynen MA (2009). Clustering of codons with rare cognate tRNAs in human genes suggests an extra level of expression regulation. *PLoS Genet.* 5: e1000548.
- Plotkin JB and Kudla G (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12: 32-42.
- Poyry TA and Jackson RJ (2011). Mechanisms governing the selection of translation initiation sites on foot-and-mouth disease virus RNA. *J. Virol.* 85: 10178-10188.
- Sanchez G, Bosch A and Pinto RM (2003). Genome variability and capsid structural constraints of hepatitis A virus. *J. Virol.* 77: 452-459.
- Sharp PM and Li WH (1986). An evolutionary perspective on synonymous codon usage in unicellular organisms. *J. Mol. Evol.* 24: 28-38.
- Smith DW (1996). Problems of translating heterologous genes in expression systems: the role of tRNA. *Biotechnol. Prog.* 12: 417-422.
- Tuller T, Carmi A, Vestsigian K, Navon S, et al. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141: 344-354.
- Wong EH, Smith DK, Rabadan R, Peiris M, et al. (2010). Codon usage bias and the evolution of influenza A viruses. Codon Usage Biases of Influenza Virus. *BMC Evol. Biol.* 10: 253.
- Wu SJ, Grouard-Vogel G, Sun W, Mascola JR, et al. (2000). Human skin Langerhans cells are targets of dengue virus infection. *Nat. Med.* 6: 816-820.
- Yadava A and Ockenhouse CF (2003). Effect of codon optimization on expression levels of a functionally folded malaria vaccine candidate in prokaryotic and eukaryotic expression systems. *Infect. Immun.* 71: 4961-4969.
- Zhou JH, Zhang J, Chen HT, Ma LN, et al. (2011). The codon usage model of the context flanking each cleavage site in the polyprotein of foot-and-mouth disease virus. *Infect. Genet. Evol.* 11: 1815-1819.
- Zhou JH, Gao ZL, Sun DJ, Ding YZ, et al. (2013a). A comparative analysis on the synonymous codon usage pattern in viral functional genes and their translational initiation region of ASFV. *Virus Genes* 46: 271-279.
- Zhou JH, Gao ZL, Zhang J, Ding YZ, et al. (2013b). The analysis of codon bias of foot-and-mouth disease virus and the adaptation of this virus to the hosts. *Infect. Genet. Evol.* 14: 105-110.