Research Article

Patterns of Synonymous Codon Usage on Human Metapneumovirus and Its Influencing Factors

Qiao Zhong, Weidong Xu, Yuanjian Wu, and Hongxing Xu

Department of Laboratory Medicine, Suzhou Municipal Hospital Affiliated Nanjing Medical University, 26 Daoqian Street, Jiangsu, Suzhou 215002, China

Correspondence should be addressed to Qiao Zhong, zhongqiao83@yahoo.com.cn

Human metapneumovirus (HMPV) is an important agent of acute respiratory tract infection in children, while its pathogenicity and molecular evolution are lacking. Herein, we firstly report the synonymous codon usage patterns of HMPV genome. The relative synonymous codon usage (RSCU) values, effective number of codon (ENC) values, nucleotide contents, and correlation analysis were performed among 17 available whole genome of HMPV, including different genotypes. All preferred codons in HMPV are ended with A/U nucleotide and exhibited a great association with its high proportion of these two nucleotides in their genomes. Mutation pressure rather than natural selection is the main influence factor that determines the bias of synonymous codon usage in HMPV. The complementary pattern of codon usage bias between HMPV and human cell was observed, and this phenomenon suggests that host cells might be also act as an important factor to affect the codon usage bias. Moreover, the codon usage biases in each HMPV genotypes are separated into different clades, which suggest that phylogenetic distance might involve in codon usage bias formation as well. These analyses of synonymous codon usage bias in HMPV provide more information for better understanding its evolution and pathogenicity.

1. Introduction

Human metapneumovirus (HMPV) is a negative single-stranded RNA virus of the family Paramyxoviridae and closely related to the avian metapneumovirus (AMPV) subgroup C [1, 2]. HMPV is an important aetiological agent of respiratory tract infection (RTI) in infants, or senior and immunocompromised individuals. This infection caused different symptoms ranging from influenza like syndromes (i.e., fever, cough, and rhinorrhea) to severe lower respiratory tract infection. Previous studies have shown that many children exposed to this virus and also easily to be reinfected as common [3–5]. Therefore, HMPV is becoming as a major concern in child respiratory tract viral infection. However, its pathogenicity is still unclear.

Genome sequencing and comparative analysis provides us a useful approach to analyze the pathogenicity of organisms. Moreover, this analysis might also provide us an approach to understand its evolution history and cell-host interaction. As previously reported HMPV genome is approximately 13 Kb in length, and the gene composition from 3′ terminal to 5′ terminal is N-P-M-F-M2-1/M2-2-SH-G-L [6, 7]. Comparative analysis suggests that its genomic organization is similar to human respiratory syncytial virus (HRSV), which just lacks 2 nonstructural genes, NS1 and NS2. Moreover, HMPV has been demonstrated the existence with two main genetic lineages termed as subtype A and B, which also containing within them the subgroups A1/A2 and B1/B2, respectively [2]. The genetic diversity analysis shows the A2 sublineage exhibits the greatest diversity among all the sublineages of HMPV.

As we all know, there are differences in the frequency of occurrence on synonymous codons in coding DNA, which termed as synonymous codon usage bias. Briefly, there are 64 different codons (61 codons encoding for amino acids plus 3 stop codons) in each organism, but only 20 different translated amino acids. These alternative codons for the same amino acids are termed as synonymous codons. In general, codon usage variation may be the product of natural selection and/or mutation pressure for accurate and efficient

translation in various organisms [8–10]. Synonymous codon usage bias on virus can provide us with a better understanding on the evolution profile, gene expression, and virus-host interaction [11–14]. However, there is still lacking about codon usage pattern of HMPV genome and its major influence factors. Herein, we firstly performed the comparative analysis of synonymous codon usage in HMPV genomes and analyzed their influencing factors. This study will provide a new insight to understand the pathogenicity and its evolution history of HMPV.

2. Materials and Methods

2.1. HMPV Genome Sequences. In this study, a total 17 complete HMPV genomes which representing two genotypes were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/) until December, 2011. The serial number (SN), Genbank number, genotype, and other information are listed in Table 1. Moreover, 10 AMPV genomes sequences were retrieved from NCBI database as reference frame (in Supplemental Table 1 supplementary material available online at doi:10.1155/2012/460837).

2.2. Analysis of Codon Usage Pattern. To investigate the characteristics of synonymous codon usage, relative synonymous codon usage (RSCU) values of each complete coding region in 17 HMPV genomes and 10 AMPV genomes were calculated [15]. The RSCU value of each codon for their amino acid was calculated as previously described [16]. A codon with an RSCU value of more than 1.0 has a positive codon usage bias, while a value of less than 1.0 has a negative codon usage bias. When the codon with RSCU values close to 1.0, it means that this codon is chosen equally and randomly. The codon usage data of human cell and bird cell were obtained from the codon usage database online (http://www.kazusa.or.jp/codon/) [17].

The effective number of codons (ENC) is used to measure deviation from expected random codon usage of HMPV and is independent of hypotheses involving natural selection [18]. The ENC values range from 20 to 61. If only one codon is used for each amino acid, this value would be 20, while all of codons are used equally, it will be 61. Moreover, the index of GC3s was used to calculate the fraction of the nucleotides G + C content at the synonymous third codon position (excluding AUG [Met], UGG [Trp], and the termination codons) [19].

2.3. Correspondence Analysis. Multivariate statistical analysis can be used to explore the relationships between variables and samples [20]. In this study, correspondence analysis was used to investigate the major trend in codon usage variation among genomes. In this study, the complete coding region of all 17 HMPV genomes was represented as a 59 dimensional vector, and each dimension corresponds to the RSCU value of one sense codon (excluding Met, Trp, and the termination codons).

TABLE 1: 17 HMPV genomes sequences used in this study.

SN	Genbank accession	Genotype	Source
1	AB503857.1	A2	Shiga, Japan
2	AF371337.2	A1	Amsterdam, The Netherlands
3	AY297748.1	B2	Quebec, Canada
4	AY297749.1	A2	Quebec, Canada
5	AY525843.1	B1	Amsterdam, The Netherlands
6	DQ843658.1	B1	Beijing, China
7	DQ843659.1	A2	Beijing, China
8	EF535506.1	B2	Taiwan, China
9	FJ168778.1	B2	Amsterdam, The Netherlands
10	FJ168779.1	A2	Amsterdam, The Netherlands
11	GQ153651.1	A2	Guangzhou, China
12	HM197719.1	B2	Rwanda
13	JN184399.1	A1	Nashville, TN, USA
14	JN184400.1	A2	Nashville, TN, USA
15	JN184401.1	B2	Nashville, TN, USA
16	JN184402.1	B2	Nashville, TN, USA
17	NC_004148.2	A2	Quebec, Canada

Abbreviation: SN is serial number.

2.4. Correlation Analysis. Correlation analysis was used to identify the relationship between nucleotide composition and synonymous codon usage pattern [21]. This analysis was implemented based on the Spearman's rank correlation analysis way.

All statistical processes were carried out with statistical software STATA11.5 for windows.

3. Results

3.1. Pattern of Synonymous Codon Usage on HMPV. In order to investigate the synonymous codon usage of HMPV, we calculated various RSCU values of various codons from 17 different strains, including different genotypes. As shown in Table 3, the preferred codons in HMPV are GCA, AGA, AAU, GAU, UGU, CAA, GAA, GGA, CAU, CAU, AUU, UUA, AAA, UUU, CCA, UCA, AGU, ACA, UAU, GUU. Interestingly, all preferred codons in HMPV genomes are ended with A/U, while none of them is ended with G/C. This result suggests that HMPV genome has a great synonymous codon usage bias, and this phenomenon might highly associate with the nucleotide composition in its genomes. Therefore, we analyzed the GC content among HMPV genomes. As shown in Table 2, the G + C content of HMPV genome is 36.91%, which shares similar extent with another RNA virus. There are over 68.57% codons are ended with A/U among HMPV genomes, and 40.87% codons are A3 end, and 27.7% codons are U3 end. This high abundance of A/U nucleotides is consistent with all preferred codons, which are ended with A/U. This phenomenon reflects that nucleotide composition is the main force to affect the codon usage bias in HMPV genome. Moreover, as shown in Table 2, the ENC values among HMPV genomes show a range from 45.127 to 48.28, and its average value of 45.785 and SD value of 0.8458.

TABLE 2: Nucleotide contents and effective number of codons in complete coding region in HMPV genomes.

SN	A%	U%	G%	C%	GC%	A3%	U3%	C3%	G3%	GC3%	*ENC
1	37.51	25.18	19.52	17.8	37.3	41.76	27.7	17.68	11.6	29.28	45.741
2	37.49	25.58	19.55	17.4	36.93	41.67	28.6	16.83	11.52	28.35	45.148
3	37.78	25.27	19.63	17.3	36.95	36.09	21.5	15.18	17.51	32.69	47.608
4	37.4	25.5	19.61	17.5	37.09	41.55	27.9	17.44	11.76	29.2	45.667
5	37.58	25.41	19.76	17.3	37.02	41.32	28.1	16.71	12.47	29.18	45.529
6	37.65	25.64	19.75	17	36.71	41.54	28.5	16.25	12.35	28.6	45.662
7	37.33	25.66	19.59	17.4	37.01	36.36	23.1	16.45	15.7	32.16	48.28
8	37.53	26.27	19.88	16.3	36.2	41.52	29.3	15.5	12.34	27.84	45.127
9	37.61	25.69	19.72	17	36.7	41.48	28.8	16.05	12.36	28.41	45.468
10	37.46	25.58	19.51	17.5	36.96	41.62	28.2	17.26	11.58	28.85	45.707
11	37.22	25.83	19.76	17.2	36.95	41.16	28.9	16.56	12.07	28.63	45.379
12	37.73	25.68	19.7	16.9	36.59	41.55	28.9	15.85	12.42	28.27	45.28
13	37.42	25.72	19.59	17.3	36.86	41.47	28.9	16.61	11.67	28.28	45.274
14	37.4	25.47	19.61	17.5	37.13	41.38	28.1	17.46	11.77	29.24	45.58
15	37.47	25.51	19.87	17.2	37.02	41.16	28.4	16.52	12.64	29.16	45.442
16	37.61	25.43	19.78	17.2	36.96	41.56	28.4	16.32	12.45	28.77	45.778
17	37.4	25.5	19.61	17.5	37.09	41.55	27.9	17.44	11.76	29.2	45.667
Mean	37.51	25.58	19.67	17.2	36.91	40.87	27.7	16.6	12.59	29.18	45.785
SD	0.144	0.241	0.114	0.33	0.25	1.755	2.1	0.713	1.589	1.296	0.8458

Abbreviation: SN is serial number; *ENC is effective number of codons.

The stable ENC values suggest that their genomic compositions are much conserved among HMPV genomes.

3.2. Nucleotide Contents of HMPV Genomes.

Natural selection and mutation pressure have been considered to be two key factors which have effect on codon usage patterns of organisms [22]. In order to investigate whether mutation pressure or natural selection as a determinative factor for codon usage mutation in HMPV, we calculated correlation relation between A%, U%, G%, C%, GC% and A3%, U3%, G3%, C3%, GC3%. As shown in Table 4, there exhibits a very complex correlation map observed in nucleotide compositions. In detail, U3% has a significant positive correlation with U% ($r = 0.7821$, $P < 0.01$), while shared negative correlation with C% ($r = -0.7153$, $P < 0.01$) and GC% ($r = -0.7474$, $P < 0.01$). C3% has positive correlation with C% ($r = 0.8331$, $P < 0.01$) and GC% ($r = 0.7880$, $P < 0.01$), and negative correlation with A% ($r = -0.6199$, $P < 0.01$) and G% ($r = -0.5846$, $P < 0.05$). GC3% has significant positive correlation with C% ($r = 0.7178$, $P < 0.01$) and GC% ($r = 0.8052$, $P < 0.01$), but has negative correlation with U% ($r = -0.7342$, $P < 0.01$). Interestingly, the highest third end nucleotide A3% has no correlation with any nucleotides. Interestingly, the GC3% shows positive correlation with C%, while shows negative correlation with G%. We calculated the GC3 skew by using formula as CG3 skew = (C3 − G3)/(C3 + G3) [22]. The GC3 skew of HMPV range s from 0.371 to −0.592, which reveals that GC composition involved in the codon usage bias. These data suggest that the nucleotide constraint might play an important role in influencing synonymous codon usage bias.

Furthermore, the correlation analysis of first two principle axes (f_1' and f_2') in HMPV and its nucleotide contents were performed (Table 5). Apparently, the first principle axis (f_1') has a significantly negative correlation with U3%, and negative correlation with GC3%. This result suggests that nucleotide U3% and GC3% are the major factor influencing the synonymous codon usage pattern in HMPV genome. Moreover, we observed the second principle axis (f_2') shared a significant positive correlation with G3%, and negative correlation with C3% and GC3%. Therefore, compositional constraint is a major factor which is involved in shaping the pattern of synonymous codon usage bias in HMPV genome.

3.3. Phylogenetic Distant Effect on Synonymous Codon Usage.

To investigate the effect of different HPMV genotypes on synonymous codon usage, we analyzed the codon usage bias of different genotypes with correspondence analysis. From the correspondence analysis, the first dimension variable f_1' and the second dimension variable f_2' can reflect 43.27% and 33.38% of total mutation, respectively. As the plot of correspondence analysis shown (Figure 1), each genotype is mainly separated and clustered into two clades. This phenomenon implied that phylogenetic relationship has some extent effect on codon usage bias. However, the sublineage of each genotype did not exhibit any significant difference among them. But this might be due to the limited number of HPMV genomes available in current study. Therefore, the phylogenetic distant might effect on the variation of synonymous codon usage in HMPV, and this difference might reflect on their biological effect, such as viral replication, virulence, and so forth.

TABLE 3: Synonymous codon usages of complete coding region of HMPV, human cell, AMPV, and bird cell.

AA	Codon	[a]RSCU	[b]RSCU	[c]RSCU	[d]RSCU
Ala	GCU	1.36	1.08	1.02	**1.42**
	GCC	0.55	**1.6**	0.77	1.33
	GCA	**1.89**	0.92	**1.99**	0.98
	GCG	0.2	0.44	0.22	0.27
Arg	CGU	0.12	0.48	0.15	0.00
	CGC	0.13	1.08	0.14	0.55
	CGA	0.31	0.66	0.34	0.27
	CGG	0.23	1.2	0.41	1.09
	AGA	**3.58**	**1.26**	**2.86**	**2.18**
	AGG	1.63	**1.26**	2.10	1.91
Asn	AAU	**1.13**	0.94	**1.08**	**1.41**
	AAC	0.87	**1.06**	0.92	0.59
Asp	GAU	**1.26**	0.92	**1.12**	0.88
	GAC	0.74	**1.08**	0.89	**1.12**
Cys	UGU	**1.04**	0.92	**1.11**	0.67
	UGC	0.96	**1.08**	0.89	**1.33**
Gln	CAA	**1.27**	0.54	**1.19**	0.48
	CAG	0.73	**1.46**	0.81	**1.52**
Glu	GAA	**1.26**	0.84	**1.11**	0.75
	GAG	0.74	**1.16**	0.89	**1.25**
Gly	GGU	1.08	0.64	0.89	0.43
	GGC	0.73	**1.36**	0.68	1.19
	GGA	**1.49**	1	**1.43**	1.08
	GGG	0.71	1	1.01	**1.30**
His	CAU	**1.13**	0.84	1.00	0.35
	CAC	0.87	**1.16**	1.00	**1.65**
Ile	AUU	**1.2**	1.08	1.02	**1.78**
	AUC	1	**1.41**	0.90	1.03
	AUA	0.8	0.51	**1.08**	0.19
Leu	UUA	**1.44**	0.48	1.19	0.42
	UUG	1.08	0.78	**1.36**	0.85
	CUU	1.14	0.78	0.83	0.56
	CUC	0.64	1.2	0.57	0.99
	CUA	0.85	0.42	0.92	0.28
	CUG	0.84	**2.4**	1.13	**2.89**
Lys	AAA	**1.21**	0.86	**1.12**	0.93
	AAG	0.79	**1.14**	0.88	**1.07**
Phe	UUU	**1.32**	0.92	**1.27**	**1.33**
	UUC	0.68	**1.08**	0.73	0.67
Pro	CCU	1.15	1.16	1.12	1.22
	CCC	0.82	**1.28**	0.74	1.22
	CCA	**1.8**	1.12	**1.78**	**1.47**
	CCG	0.23	0.44	0.36	0.08
Ser	UCU	0.97	1.14	0.93	0.71
	UCC	0.63	1.32	0.72	1.15
	UCA	**1.46**	0.9	**1.37**	1.06
	UCG	0.13	0.3	0.18	0.26
	AGU	**1.46**	0.9	**1.41**	0.44
	AGC	1.36	**1.44**	1.30	**2.38**

TABLE 3: Continued.

AA	Codon	[a]RSCU	[b]RSCU	[c]RSCU	[d]RSCU
Thr	ACU	1.17	1	0.98	0.16
	ACC	0.92	**1.44**	0.83	1.28
	ACA	**1.75**	1.12	**1.91**	**1.76**
	ACG	0.16	0.44	0.27	0.80
Tyr	UAU	**1.34**	0.88	**1.14**	0.95
	UAC	0.66	**1.12**	0.86	**1.05**
Val	GUU	**1.46**	0.72	1.17	0.57
	GUC	0.73	0.96	0.72	0.91
	GUA	0.93	0.48	0.87	0.80
	GUG	0.88	**1.84**	**1.24**	**1.71**

The number in this table is shown as relative synonymous codon usage values (RSCU) and [a]RSCU values represent values in HMPV, [b]RSCU values represent values in human host cell, [c]RSCU values represent values in AMPV, and [d]RSCU values represent values in bird cell. All preferred codons are shown in bold. AA is abbreviation of amino acid.

TABLE 4: Summary of correlation analysis between A%, U%, C%, G%, and GC% and A3, U3, C3, G3, and GC3 in all HMPV samples.

	A3%	U3%	C3%	G3%	GC3%
A%	0.0959[NS]	0.1144[NS]	−0.6199**	0.3788[NS]	−0.2595[NS]
U%	−0.1596[NS]	0.7821**	−0.4150[NS]	−0.0933[NS]	−0.7342**
C%	0.2417[NS]	−0.7153**	0.8331**	−0.4699[NS]	0.7178**
G%	−0.4012[NS]	0.4234[NS]	−0.5846*	0.6129**	−0.3138[NS]
GC%	0.0774[NS]	−0.7474**	0.7880**	−0.1328[NS]	0.8052**

The number in the table represents as correlation coefficient r value, which is calculated from each correlation analysis. Abbreviation: NS represents as nonsignificant ($P > 0.05$), * represents $P < 0.05$, ** represents $P < 0.01$.

TABLE 5: Summary of correlation between the first two principle axes and nucleotide contents in samples.

Base compositions	f_1' (43.27%)	f_2' (33.38%)
A3%	$r = -0.2247^{NS}$	$r = -0.2491^{NS}$
U3%	$r = -0.8091^{**}$	$r = 0.4454^{NS}$
C3%	$r = 0.3106^{NS}$	$r = -0.8233^{**}$
G3%	$r = 0.1449^{NS}$	$r = 0.6393^{**}$
GC3%	$r = 0.8447^{**}$	$r = -0.5043^{*}$

The number in the table represents as correlation coefficient r value, which is calculated from each correlation analysis. Abbreviation: NS represents as nonsignificant ($P > 0.05$), * represents $P < 0.05$, ** represents $P < 0.01$.

3.4. Relationship between Codon Usage Pattern of HMPV and Its Host. From the ENC-GC3% plot analysis (Figure 2), the plots of each HMPV genomes are all under the expected curve, none of them shows above the curve. This result implied that mutation pressure is the major factor influencing the codon usage [19]. Moreover, there are still some other factors that can effect on the codon usage bias of HMPV. In the current study, we compared the patterns of codon usage in HMPV and human host. As shown in Table 3, the pattern of synonymous codon usage in HMPV shows a complementary profile, which shows in human cell. In detail,
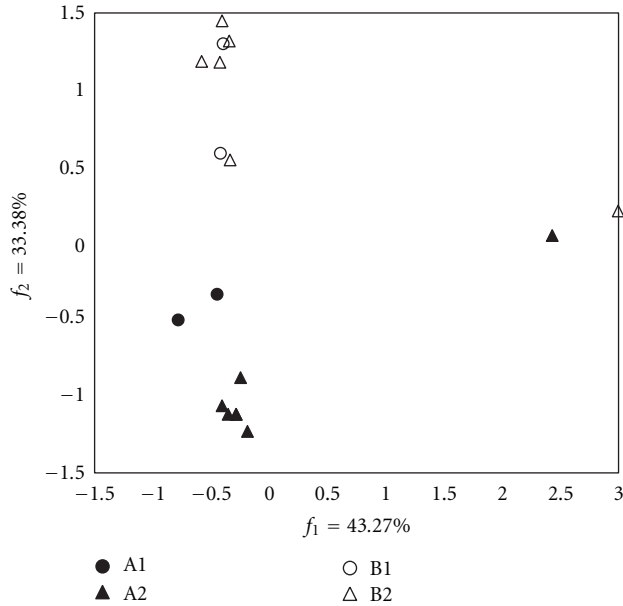
FIGURE 1: The plot of value of the first and second axis of each complete coding region in COA. The first axis (f_1') accounts for 43.27% of the total variation, and the second axis (f_2') accounts for 33.38% of the total variation. Each HMPV complete coding region was divided by genotype.
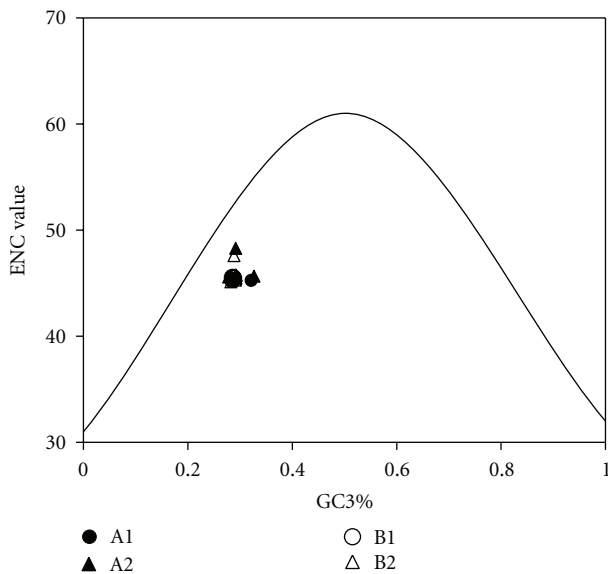

FIGURE 2: Effective number of codons used in each complete coding region plotted against the GC3. The continuous curve plots the relationship between GC3 and ENC in the absence of selection. All of plots lie below the expected curve.

HMPV and human host cell shared only 1 preferred codon (AGA), which encoded for Arginine, while there are 17 different preferred codons between them. As a reference frame, AMPVs were enrolled in this study, and it also shows a complementary pattern with its host, bird cell. The comparative analysis among HMPV and human host cell, HMPV and AMPV host bird cell, and HMPV and AMPV were analyzed. To compare the complementary ability of HMPV with bird cell, there are more overlays (6 preferred codons overlay) than HMPV with its human host cell (only 1 pre-ferred codons overlay). This result shows human cell has much more complementary pattern with HMPV than bird cells. This might be more suitable for HMPV survive and persist infect in human host environment. This result also suggests that host factor plays an important role in codon usage bias in HMPV. This complementary trend will benefit for virus replication instead of competitive with its host and it might help us to understand the mechanism of HMPV persistent infection. Interestingly, HMPVs are shares with more than 15 preferred codons to AMPV, which might be due to a close phylogenetic distance.

4. Discussion

Synonymous codon usage analysis can reveal much about virus genome. To understand the extent and causes of codon usage bias is essential for studying the viral evolution, particularly the interaction between viruses and host immune response. In this study, we analyzed the codon usage bias in HMPV, and its influencing factors. As we know, the variation and evolution of virus generally happened in the changes of nucleotide composition [23]. Therefore, the nucleotide composition bias is the main force to influence the synonymous codon usage patterns. In this study, several evidences can support this statement in HMPV genomes. First of all, in HMPV, all preferred codons are ended in A/U nucleotide, which occupied the majority of nucleotide composition in HMPV genome. This phenomenon confirmed that nucleotide composition was the main force in shaping the pattern of codon usage. Secondly, ENC was used to quantify the codon usage bias, which is one of the best overall estimators of absolute synonymous codon usage bias [18]. In this study, we observed ENC of these genomes fluctuated from 45.13 to 48.28 with a mean 45.78 ± 0.85. This ENC value of HMPV is consistent with other previously reported RNA virus in the same family Paramyxoviridae, ranging from 43.8 to 55.1, that is, Measles virus 55.1, Mumps virus 54.3, Parainfluenza-3 virus 43.8, and Respiratory syncytial virus 44.3 [8, 24, 25]. Moreover, the ENC of HMPV is more close to respiratory infection agents, RSV and parainfluenza-3, which reveals that the similar extent of codon usage bias among viruses might have similar infection syndrome. This observation helps us to address an interesting assumption that synonymous codon usage bias of virus might associate with its pathogenicity.

Mutation pressure and natural selection are generally treated as the main factors that account for codon usage bias in different organisms [22]. ENC-GC plot was considered as a part of the general strategy to investigate patterns of synonymous codon usage [9, 18, 26]. Herein, all the plots are laid below the expected curve, suggesting that codon usage bias in all these 17 HMPV genomes was principally influenced by mutation bias, which consistent with that mutation pressure rather than natural selection is the most important

determinant of the codon usage in human RNA virus [8, 27–31]. This observation can be explained as the mutation rates in RNA viruses much higher than those in DNA viruses.

Mutation pressure is the main force in shaping synonymous codon usage bias of RNA virus. However, based on correspondence analysis, we observed an interesting phenomenon that codon usage bias in HMPV showed distinct differences among different phylogenetic types. It might suggest that codon usage bias plays an important role in HMPV evolution history. This similar phenomenon also observed in several other viruses, this might reflect that phylogenetic difference is a common influencing factor in shaping codon usage bias [25, 27, 30–34]. This difference could potentially affect the viral protein expression rate or its replication manner. Therefore, we hypothesize that the difference of codon usage bias might influence its virulence in different genotypes.

In this study, we also observed that HMPV showed a complementary trend with human cells by comparing the codon usage. This complementary will be benefit for the survive of virus, which can keep replication by using the nonpreferred codons in the host cell without competition, and this could be one of the mechanisms of virus persistent infection in the human environment. Moreover, this pattern might also be caused by the longitude selection and evolution between human hosts with virus. Therefore, this characteristic is important for HMPV keeping balanced with their host on the codon usage side, and also for understanding the cell-host interaction and viral evolution.

5. Conclusions

In summary, we firstly reported the synonymous codon usage pattern in HMPV genomes and revealed that mutation pressure is the main force in shaping its codon usage bias. Phylogenetic difference and host factors are also discussed, and this information can provide better understanding on the molecular evolution and its pathogenicity of HMPV.

Acknowledgment

References

[1] S. Broor, P. Bharaj, and H. S. Chahar, "Human metapneumovirus: a new respiratory pathogen," *Journal of Biosciences*, vol. 33, no. 4, pp. 483–493, 2008.

[2] F. Feuillet, B. Lina, M. Rosa-Calatrava, and G. Boivin, "Ten years of human metapneumovirus research," *Journal of Clinical Virology*, vol. 53, no. 2, pp. 97–105, 2012.

[3] Y. Abed and G. Boivin, "Human metapneumovirus infection in immunocompromised child," *Emerging Infectious Diseases*, vol. 14, no. 5, pp. 854–856, 2008.

[4] M. Arabpour, A. Samarbafzadeh, M. Makvandi et al., "The highest prevalence of human metapneumovirus in Ahwaz children accompanied by acute respiratory infections," *Indian Journal of Medical Microbiology*, vol. 26, no. 2, pp. 123–126, 2008.

[5] M. J. Carr, A. Waters, F. Fenwick, G. L. Toms, W. W. Hall, and E. O'Kelly, "Molecular epidemiology of human metapneumovirus in Ireland," *Journal of Medical Virology*, vol. 80, no. 3, pp. 510–516, 2008.

[6] R. Piyaratna, S. J. Tollefson, and J. V. Williams, "Genomic analysis of four human metapneumovirus prototypes," *Virus Research*, vol. 160, no. 1-2, pp. 200–205, 2011.

[7] B. G. van den Hoogen, T. M. Bestebroer, A. D. M. E. Osterhaus, and R. A. M. Fouchier, "Analysis of the genomic sequence of a human metapneumovirus," *Virology*, vol. 295, no. 1, pp. 119–132, 2002.

[8] G. M. Jenkins and E. C. Holmes, "The extent of codon usage bias in human RNA viruses and its evolutionary origin," *Virus Research*, vol. 92, no. 1, pp. 1–7, 2003.

[9] G. A. Palidwor, T. J. Perkins, and X. Xia, "A general model of Codon bias due to GC mutational bias," *PLoS ONE*, vol. 5, no. 10, Article ID e13431, 2010.

[10] P. M. Sharp, L. R. Emery, and K. Zeng, "Forces that influence the evolution of codon bias," *Philosophical Transactions of the Royal Society B*, vol. 365, no. 1544, pp. 1203–1212, 2010.

[11] T. Warnecke and L. D. Hurst, "Evidence for a trade-off between translational efficiency and splicing regulation in determining synonymous codon usage in *Drosophila melanogaster*," *Molecular Biology and Evolution*, vol. 24, no. 12, pp. 2755–2762, 2007.

[12] N. Stoletzki and A. Eyre-Walker, "Synonymous codon usage in *Escherichia coli*: selection for translational accuracy," *Molecular Biology and Evolution*, vol. 24, no. 2, pp. 374–381, 2007.

[13] A. O. Urrutia and L. D. Hurst, "Codon usage bias covaries with expression breadth and the rate of synonymous evolution in humans, but this is not evidence for selection," *Genetics*, vol. 159, no. 3, pp. 1191–1199, 2001.

[14] Y. Kim, "Effect of strong directional selection on weakly selected mutations at linked sites: implication for synonymous codon usage," *Molecular Biology and Evolution*, vol. 21, no. 2, pp. 286–294, 2004.

[15] S. G. E. Andersson and P. M. Sharp, "Codon usage and base composition in *Rickettsia prowazekii*," *Journal of Molecular Evolution*, vol. 42, no. 5, pp. 525–536, 1996.

[16] M. Bulmer, "The selection-mutation-drift theory of synonymous codon usage," *Genetics*, vol. 129, no. 3, pp. 897–907, 1991.

[17] Y. Nakamura, T. Gojobori, and T. Ikemura, "Codon usage tabulated from international DNA sequence databases: status for the year 2000," *Nucleic Acids Research*, vol. 28, no. 1, article 292, 2000.

[18] A. Fuglsang, "The effective number of codons for individual amino acids: some codons are more optimal than others," *Gene*, vol. 320, no. 1-2, pp. 185–190, 2003.

[19] E. Elhaik, G. Landan, and D. Graur, "Can GC content at third-codon positions be used as a proxy for isochore composition?" *Molecular Biology and Evolution*, vol. 26, no. 8, pp. 1829–1833, 2009.

[20] M. V. Boland and R. F. Murphy, "Multivariate analysis," in *Current Protocols in Cytometry*, chapter 10, p. unit 108, 2001.

[21] C. B. Bagwell, "Data analysis through modeling," in *Current Protocols in Cytometry*, chapter 10, p. unit 107, 2001.

[22] T. V. Tatarinova, N. N. Alexandrov, J. B. Bouck, and K. A. Feldmann, "GC3 biology in corn, rice, sorghum and other grasses," *BMC Genomics*, vol. 11, no. 1, article 308, 2010.

[23] G. Sablok, K. C. Nayak, F. Vazquez, and T. V. Tatarinova, "Synonymous codon usage, GC3, and evolutionary patterns across plastomes of three pooid model species: emerging grass

genome models for monocots," *Molecular Biotechnology*, vol. 49, no. 2, pp. 116–128, 2011.

[24] T. Zhou, W. Gu, J. Ma, X. Sun, and Z. Lu, "Analysis of synonymous codon usage in H5N1 virus and other influenza A viruses," *BioSystems*, vol. 81, no. 1, pp. 77–86, 2005.

[25] W. Q. Liu, J. Zhang, Y. Q. Zhang et al., "Compare the differences of synonymous codon usage between the two species within cardiovirus," *Virology Journal*, vol. 8, article 325, 2011.

[26] J. M. Comeron and M. Aguadé, "An evaluation of measures of synonymous codon usage bias," *Journal of Molecular Evolution*, vol. 47, no. 3, pp. 268–274, 1998.

[27] X. S. Liu, Y. G. Zhang, Y. Z. Fang, and Y. L. Wang, "Patterns and influencing factor of synonymous codon usage in porcine circovirus," *Virology Journal*, vol. 9, article 68, 2012.

[28] M. R. Ma, X. Q. Ha, H. Ling et al., "The characteristics of the synonymous codon usage in hepatitis B virus and the effects of host on the virus in codon usage pattern," *Virology Journal*, vol. 8, Article ID Article number544, 2011.

[29] Y. S. Liu, J. H. Zhou, H. T. Chen et al., "The characteristics of the synonymous codon usage in enterovirus 71 virus and the effects of host on the virus in codon usage pattern," *Infection, Genetics and Evolution*, vol. 11, no. 5, pp. 1168–1173, 2011.

[30] J. H. Zhou, J. Zhang, H. T. Chen, L. N. Ma, and Y. S. Liu, "Analysis of synonymous codon usage in foot-and-mouth disease virus," *Veterinary Research Communications*, vol. 34, no. 4, pp. 393–404, 2010.

[31] P. Tao, L. Dai, M. Luo, F. Tang, P. Tien, and Z. Pan, "Analysis of synonymous codon usage in classical swine fever virus," *Virus Genes*, vol. 38, no. 1, pp. 104–112, 2009.

[32] Y. S. Liu, J. H. Zhou, H. T. Chen et al., "Analysis of synonymous codon usage in porcine reproductive and respiratory syndrome virus," *Infection, Genetics and Evolution*, vol. 10, no. 6, pp. 797–803, 2010.

[33] S. Zhao, Q. Zhang, X. Liu et al., "Analysis of synonymous codon usage in 11 Human Bocavirus isolates," *BioSystems*, vol. 92, no. 3, pp. 207–214, 2008.

[34] P. L. Meintjes and A. G. Rodrigo, "Evolution of relative synonymous codon usage in human immunodeficiency virus type-1," *Journal of Bioinformatics and Computational Biology*, vol. 3, no. 1, pp. 157–168, 2005.