Acta virologica **66**: 309 – 323, 2022 doi: 10.4149/av_2022_402

Insights into the factors affecting synonymous codon usage bias in the coat protein of soil-borne sugar beet-infecting viruses

Shirin Farzadfar, Reza Pourrahim

Plant Virus Research Department, Iranian Research Institute of Plant Protection (IRIPP), Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran

Received February 14, 2022; revised July 25, 2022; accepted October 7, 2022

Summary. – Synonymous codon usage patterns (CUP) indicate a series of evolutionary changes that influence viral survival rates and fitness. As viral coat protein (*CP*) genes evolve more rapidly, they provide a strong incentive to study the genetic diversity and evolutionary changes of viruses based on *CP* genes. In this study, CP sequences were analyzed using different bioinformatic approaches. Nucleotide composition and relative synonymous codon usage (RSCU) analysis indicates mutation bias and prefers A/U ended codon. The Parity rule 2 (PR2) and the effective number of codon (ENC) plots showed the dominant role of mutation in CUP of *CP* genes. Host relatedness in terms of codon usage was analyzed, with codon adaptation index (CAI), and similarity index (SiD) calculations; indicating that during the evolution of these viruses, sugar beet has a greater impact on the beet black scorch virus (BBSV) rather than beet necrotic yellow vein virus (BNYVV). In addition, *Polymyxa betae* has a more profound effect on shaping RSCU patterns of BNYVV. As *P. betae* and *Olpidium brassica* are the natural reservoir and host for soil-borne viruses infecting sugar beet, it makes sense that the virus has evolved its genomic properties to a steady level to better adjust to its primary host's condition. The principal component (PCA) plot indicated that BNYVV and beet soil-borne virus (BSBV) have originated from the Old World, whereas BBSV separated into three diverged groups, what is in accordance with phylogenetic groups. This research is the first study of codon usage analysis of beet soil-borne viruses and increases our knowledge about the mechanisms that support codon usage and their evolution.

Keywords: beet soil-borne viruses; codon usage patterns; mutation pressure; natural selection; host adaptation

Introduction

Among the wide range of organisms that threaten the productivity of beets (*Beta vulgaris* L.), soil-borne viruses are economically considerable due to yield losses worldwide (Henry, 1996; Whitney and Duffus, 1998; Biancardi and Tamada, 2016). Rhizomania caused by beet necrotic yellow vein virus (BNYVV; the genus *Benyvirus*; the *Benyviridae* family) is one of the most destructive diseases and constrains sugar beet production (up to 80%) with widespread distribution in the beet cultivated areas of the world (Biancardi and Tamada, 2016) since it was first reported in Italy in 1950s (Canova, 1959). BNYVV was then reported from Japan (1965) (Tamada and Baba, 1973), and in Nei Menggu (Inner Mongolia) from China in 1978 (Gao *et al.*, 1983). During the 1970s to 1980s BNYVV was reported from European countries (Putz and Vuittenez, 1974; Hamdorf *et al.*, 1977; Kouyeas, 1979; Hill and Torrance, 1989; Suárez *et al.*, 1999; Borodynko *et al.*, 2009). BNYVV spread from central and southern Europe to northern and eastern Europe and

E-mail: farzadfar2002@yahoo.com; phone: +98-21-22403012.

Abbreviations: BBSV = beet black scorch virus; BNYVV = beet necrotic yellow vein virus; BSBV = beet soil-borne virus; BVQ = beet virus Q; CAI = codon adaptation index; CP = coat protein; ENC = effective number of codons; PC = Parity rule 2; PCA = principal component; RSCU = relative synonymous codon usage; SiD = similarity index

the Middle East during a few decades (Asher, 1993; Tamada, 1999; McGrann *et al.*, 2009). In the United States, the disease was first reported in California in 1983 (Duffus *et al.*, 1984). BNYVV has been reported from Iran in 1996 (Izadpanah *et al.*, 1996). Two soil-borne sugar beet pomoviruses, beet soil-borne virus (BSBV) and beet virus Q (BVQ) (the genus *Pomovirus*; the *Virgaviridae* family) (Adams *et al.*, 2017) are frequently detected in sugar beet samples infected with BNYVV (Meunier *et al.*, 2003; Borodynko *et al.*, 2009). The coat protein (CP) gene of these three viruses is located on RNA2 (Adams *et al.*, 2017) with a common vector "a soil-borne protoctist, *Polymyxa betae*" (Cui 1988; Tamada, 2002). Beet black scorch virus (BBSV) from the genus *Betanecrovirus*, is another sugar beet infecting soil-borne virus. BBSV belongs to the family *Tombusviridae*, has a chytrid vector *Olpidium brassica* (Zhang *et al.*, 1996) and it was reported from China, some European countries, Iran, and the USA (Zhang *et al.*, 1996; Weiland *et al.*, 2007; Koenig and Valizadeh, 2008; Farzadfar and Pourrahim, 2019).

Understanding the evolution of virus-host interactions is important, due to rapid evolution through genetic recombination, mutation, the potential of adaption to new or resistant hosts (Garcia-Arenal *et al.*, 2001; Davino *et al.*, 2017), fast adaptation to the different environmental conditions, and mostly lack of effective chemical compounds (Elena *et al.*, 2014). Identification of codon usage patterns provides important information about the host-pathogen co-evolution, molecular evolution of genes, and adaptation of the virus to a specific host. Generally, composition-biased selection pressure and/or mutation pressure for accurate and efficient translation in various organisms are the main reasons for this bias. Any similarities in the codon usage pattern result in some degrees of biological relationship, environmental adaptation, and evolution. The recent advancements in sequencing technologies, allow studying the codon usage behavior of viral diseases (Xu *et al.*, 2008; Liu *et al.*, 2012; He *et al.*, 2017; He *et al.*, 2019; Gu *et al.*, 2020).

It is presumed that viral coat protein (*CP*) genes evolved more rapidly than proteins involved in the replication and expression of virus genomes (Callaway *et al.*, 2001), thus providing a strong incentive to study the diversity of viruses based on *CP* genes. The genetic variability and population structure of beet soil-borne viruses have already been studied (Chiba *et al.*, 2011; Biancardi and Tamada, 2016). We determined CP sequences of new Iranian isolates of BNYVV (n = 4), BBSV (n = 5), and BSBV (n = 6). Then the genotyping profile based on the comparison of the nucleotide sequences of the *CP* gene with other isolates available in GenBank was analyzed using different bioinformatic approaches. In addition, the codon usage patterns, among these viruses have been characterized. Understanding the synonymous codon usage can provide significant insights into the evolution, host adaptation, virus pathogenesis, and survival. These analyses reveal novel information about the evolutionary fitness of beet soil-borne viruses.

Materials and Methods

Viral isolates. Soil samples from 21 sugar beet fields of Kermanshah and Hamadan provinces in West Iran were collected during years 2018–2019. Soil samples were assayed for BNYVV, BSBV, BVQ, and BBSV in the greenhouse (Table 1). An autoclaved potting soil was used as control. The susceptible sugar beet cultivar (Jolge) was planted in the soil samples that were mixed with equal parts of autoclaved sand. One month after planting, the plants were removed (five plants for each soil sample) and tip roots were used for enzyme-linked immunosorbent assay (ELISA) (Clark and Adams, 1977) by specific antibodies to BNYVV (Bioreba, Switzerland), BSBV (As-0576.1), BSBV/BVQ (As-0576.2) (DSMZ, Germany), and for BBSV LOEWE (Sauerlach, Germany). In summary, microtitre plates (Maxisorb, NUNC, Denmark) were coated with the recommended titer of IgG in carbonate buffer. Root tip tissues were ground in extraction buffer (1:5 wt/vol); then added to the plate wells and incubated overnight at 4°C. Absorbance values were measured at 405 nm, using a BioTek microplate reader 1 h after the addition of the *p*-nitrophenyl phosphate substrate. Samples with OD405 values equal to or more than three mean of the negative controls were considered positive.

Reverse transcription-polymerase chain reaction (RT-PCR). Total RNAs were extracted from ELISA positive samples of infected sugar beet using RNeasy plant mini kit (QIAGEN, USA) according to the instructions of the kit manufacturer. Complete *CP* genes were amplified using specific primers designed in this study (Table 2). Each cDNA synthesis reaction contained 4 µl of 5x RT buffer (Sinaclon, Iran), 1 µl (200 U) of M-MuLV reverse

Table 1. Occurrence of beet soil-borne viruses in soil samples collected from Kermanshah and Hamadan provinces

	No. of fields/root samples tested	Number of infected sample[a] (%)[b]		
County/Yare		BNYVV	BSBV	BBSV
Kermanshah/2018	12 / 60	41[a] (68.33)	26 (43.33)	12 (20.00)
Hamadan/2019	9 / 45	35 (77.70)	14 (31.11)	8 (17.78)
Total	21 / 105	76 (72.38)	40 (38.05)	20 (19.05)

[a]Based on serological reactions (ELISA). [b]The percentage of viral infection.

transcriptase (Sinaclon), 2 µl of 10 mM dNTPs mix, 0.5 µl (20 U) ribonuclease inhibitor (Vivantis, Malaysia), 1 µl of each reverse primer (10 pmol/µl), 7.5 µl nuclease-free water, and 5 µl of total RNA extraction with final volume 20 µl. The microtubes were incubated at 42°C for 1 h and at 70°C for 10 min to inactivate the M-MuLV reverse transcriptase. PCR reactions were amplified in 50 µl final volume, containing 4 µl of cDNA product, 5 µl of 10x PCR buffer (Sinaclon), 1 µ of 50 mM $MgCl_2$, 1 µl of 10 mM dNTP mix, 1 µl (20 pmol) of each forward and reverse specific primers, 1 µl (5 U) of *Pfu* DNA polymerase (Sinaclon). PCR program was 94°C for 2 min; 30 cycles of 94°C for 30 s, 58°C for the 30 s, and 72°C for 30 s, followed by 72°C for 10 min. PCR products were evaluated by electrophoresis in a 1% agarose gel in TBE buffer including ethidium bromide (final concentration 1 µg/ml). The PCR products were isolated from agarose gel and afterwards purified using Wizard PCR DNA purification kit (Promega) according to the manufacturer's recommendations. Purified DNA fragments were cloned into plasmid pGEM-T (Promega) according to the manufacturer's recommendations. PCR products or cloned fragments were sequenced in both directions and three replications using commercial service. Sequenced data were assembled using BIOEDIT version 5.0.9 (Hall, 1999). The sequence was analyzed using the BLAST tool in the National Center for Biotechnology Information (NCBI).

Nucleotide composition analysis of coat protein. After deleting five non-bias codons including AUG (start codon), UGG (encoding Trp), and three termination codons UAA, UGA, and UAG, the component parameters of the CP sequences were calculated. The total percent nucleotide composition and the overall GC and AU contents were estimated by MEGAX (Kumar *et al.*, 2018). The CodonW 1.4.2 package was used for assessing the nucleotide mixtures at the 3rd codon position (A3, C3, T3, and G3%). The Emboss explorer (http://www.bioinformatics.nl/emboss-explorer/) was used for calculating GC content at the first, second, and third codon positions (GC1s, GC2s, GC3s), where the average of GC1 and GC2s is indicated by GC1,2s.

Analysis of relative synonymous codon usage (RSCU). The relative synonymous codon usage (RSCU) values for all of the CP coding sequences of BNVYY, BSBV, BVQ, and BBSV were determined to find the aspects of synonymous codon usage without the complicated effect of amino acid composition and coding sequence size of *CP* gene samples according to the previously described method (He *et al.*, 2019). The RSCU values were calculated as follows:

$$RSCU_{ij} = \frac{g_{ij}}{\sum_j^{n_i} g_{ij}} \times n_i$$

where g_{ij} corresponds to the synonymous codon usage value of the i_{th} codon for the j_{th} amino acid, and n_i is the number of synonymous codons that encode the j_{th} amino acid (Sharp and Li, 1986a).

Considering that all codons for the particular amino acid are used equally, RSCU values indicate the ratio between the observed usage prevalence of one codon in a gene sample and the expected usage frequency in the synonymous codon family. When the RSCU value is 1.0, it means there is no codon usage bias for that amino acid and the codons are chosen equally or randomly ((Sharp and Li, 1986b). The synonymous codons with RSCU values < 1.0 have negative codon usage bias and were considered as weak or less-abundant codons, whereas, synonymous codons with RSCU values > 1.0 have positive codon usage bias and were identified as abundant codons. In addition, the synonymous codons with RSCU values > 1.6 and < 0.6 were indicated as "overrepresented" and "underrepresented", respectively (Wong *et al.*, 2010).

Table 2. Primers used in this study

Virus	Primer	Expected DNA bands
BNYVV.CP-F	5' tatcaaatgttaaggcaatcgag 3'	621 bp
BNYVV CP-R	5' gagtgaaggtagatatatgacatgg 3'	
BSBV CP-F	5' atggttgatccgcggtatg 3'	530 bp
BSBV CP-R	5' cacgtacgtccactttaacc 3'	
BBSV CP-F	5' tagtatag(t/a)(c/t)aataaatggcac 3'	737 bp
BBSV CP-R	5' cccacatcctggtgtggttaat 3'	
BVQ CP-F	5' gtctagaagtatggttgatccaag 3'	522 bp
BVQ CP-R	5' caggacaattgattgctatgagcc 3'	

The effective number of codons (ENC) and ENC-plot analysis. ENC analysis was used to quantify the absolute codon usage bias by evaluating the degree of codon usage bias exhibited by the beet soil-borne viruses *CP* genes, regardless of gene length and the number of amino acids. The calculation value of ENC is as follows:

$$ENC = 2 + \frac{9}{F_2} + \frac{1}{F_3} + \frac{5}{F_4} + \frac{3}{F_6}$$

where the average values for F_k is referred to as $\bar{F_k}$(k = 2, 3, 4, 6) and k-fold degenerate amino acids are indicated by *k*. F_k is calculated with the following formulae:

$$F_k = \frac{nS - 1}{n - 1}$$

In this formula "*n*" is the overall incidences number of the codon for that amino acid.

$$S = \sum_{i=1}^{k} (\frac{n_i}{n})^2$$

Where n_i is the overall number of the i_{th} codon for the matching amino acid.

ENC value ranged from 20 (a high codon usage bias using only one of the possible synonymous codons for the matching

amino acid) to 61 (no bias using all available synonymous codons equally for the matching amino acid). Generally, the lower ENC value reflects a highly expressed gene with a great codon preference, whereas, those genes with low expression have more rare codons and show higher ENC values (Sharp *et al.*, 1986; Comeron and Aguade, 1998; Wright, 1990). Using the ENC values versus GC3s value plots (ENC-plot analyses), the effect of mutational pressure on codon usage bias was analyzed. When the points are located on the standard curve it shows that mutation pressure is the only factor for driving the codon usage bias. Otherwise, the effect of other factors like natural selection must be considered. Using the following formula, the ENC value was determined demonstrating the value of GC3s:

$$ENC_{\text{expectation}} = 2 + s + \frac{29}{s^2 + (1-s)^2}$$

Principal component (PCA) and Parity rule 2 (PR2) analysis. PCA is a multivariate statistical method that is used to explore the relationship between variables and samples (Sharp *et al.*, 1986; Sueoka, 1999). PCA was used to analyze the major trends in codon usage patterns among beet soil-borne viruses *CP* gene sequences. Parity rule 2 (PR2) plot analysis shows the influence of natural selection and mutation pressure on the codon usage of each gene using A3/(A3+U3) plotted value versus G3/(G3+C3) value. The center of the PR2 plot is 0.5 which indicates A=U and G=C (Sueoka, 1999). If there is no deviance between mutation pressure and selection pressure, the points are placed in the center of the plot.

Influence of overall host codon usage on beet soil-borne viruses. The influence of the codon usage bias of the host (*B. vulgaris*) and vector (*P. betae*) are measured by similarity index (SiD) value (Zhou *et al.*, 2013). The SiD was determined by:

$$R(A.B) = \frac{\sum_{i=1}^{59} a_i b_i}{\sqrt{\sum_{i=1}^{59} a_i^2 \times \sum_{i=1}^{59} b_i^2}}$$

$$\text{SiD} = \frac{1 - R(A.B)}{2}$$

In this formula a_i shows the RSCU value of 59 synonymous codons of the beet soil-borne *CP* gene sequences, and b_i shows the RSCU value of the identical codons of the potential host. The SiD values vary between 0.0 to 1.0 and the higher values show that the host has a deeper effect on the usage of codons.

Codon adaptation index (CAI) analysis. CAI analysis is a quantitative method that foresees the expression level of a gene-based on its coding sequence. CAI values for beet soil-borne CP sequences were determined using the CAIcal SERVER (http://genomes.urv.cat/CAIcal/RCDI/). The CAI values ranging from 0.0 to 1.0 indicate the various degrees of adaptation to the host. The high CAI value of a sequence shows its stronger adaptability to the host, and sequences with higher CAIs are considered preferred over those with lower CAIs (Puigbo *et al.*, 2008).

Results

Beet soil-borne virus isolates

A total of 105 beetroot tip samples were tested by DAS-ELISA from which 76 (72.38%), 40 (38.05%), and 20 (19.05%) samples were found positive for BNYVV, BSBV, and BBSV viruses, respectively. The percentage of viral infection varied among collection regions. The highest percent of beet soil-borne viruses were found in Kermanshah province with 68.33%, 43.33%, and 20.00%, for BNYVV, BSBV, and BBSV, respectively. No positive samples were detected for BVQ (Table 1). After polymerase chain reaction (PCR), different expected amplicons with about 621, 530, and 737 bps were amplified for BNYVV, BSBV, and BBSV, respectively. The obtained full-length CP sequences were deposited in GenBank under Acc. Nos. OK669130 to OK669134 for BBSV, OK669135 to OK669138 for BNYVV, and OK669139 to OK669144 for BSBV.

Beet soil-borne viruses are enriched with A and U nucleotides

The potential influence of nucleotide compositional constraints on codon usage was analyzed for each beet soil-borne virus population. It was found that the nucleotides T(U), G, A, and C were the most abundant in the CP coding sequences of BNVYY, BSBV, BBSV, and BVQ with mean values of 28.92% ± 1.24%, 27.00% ± 1.01, 24.83% ± 1.31, and 19.24% ± 1.87, respectively. The mean values of AU% were 56.63 ± 0.66 for BNYVV, followed by 54.42 ± 0.53, 53.85 ± 0.09, and 51.09 ± 1.38 for BSBV, BVQ, and BBSV, respectively. The analysis of nucleotide composition at third positions of synonymous codons indicated that GC3s values ranged from 31.3 to 38.3 with an average of 35.45 ± 1.96, 42.1 to 43.9 with an average of 42.85 ± 0.92, 42.2 to 45.8 with an average 45.62 ± 0.30, and 42.1 to 44.6 with an average 42.85 ± 3.91 for BNYVV, BSBV, BQV, and BBSV, respectively.

Substantial variation in codon usage of beet soil-borne virus CP gene sequences

Total variation in codon usage was assessed by RSCU analysis, which indicates that the beet soil-borne CP coding sequences were AU-rich, and A/U ending codons were favored in the *CP* genes. Ten out of 18 frequently used codons were U-ended while 4, 2, and 2 out of 18 were ended with the A, C, and G (Table 3). Irrespective of the soil-borne viruses the RSCU values ≥ 1.6 were detected for four of the

optimal synonymous codons (UAA, GUU, ACU, and GCU), with the highest preferred value for GCU codon (1.90). The under-represented (RSCU ≤ 0.6) and non-preferred codons were mostly G/C-ended (Table 3). In addition, the variation of the RSCU values of *CP* genes were computed for each codon of beet soil-borne viruses and the results showed in three main clusters of codons (Fig. 1a-1d). The first cluster included under-represented codons (RSCU < 1), which contained A/U or G/C ending codons. The second cluster consisted of A/U-ending codons with RSCU ≥ 1. The last group consisted of G/C ending codons with RSCU ≥ 1. As indicated for BNYVV (Fig. 1a), 31 out of 59 codons ending with A/U or G/C, showed the RSCU < 1, whereas the over-represented codons (RSCU > 1.6) ending with A/U, and C were 12 and 1, respectively (Fig. 1a). Some differences were indicated among A and B type populations (e.g., the two codons UAU and UAC for tyrosine amino acid). The RSCU based on each BSBV isolate shows 30 out of 59 codons ending with A/U or G/C with RSCU < 1, and those of them with RSCU > 1.6 are A/U ending codons (4 out of 59) with a higher RSCU value for the UUU codon, and G/C ending codons (3 out of 59) (Fig. 1b). For the *CP* gene of BVQ the most RSCU > 1.6 were detected among A/U and G ending codons (7 and 4), respectively (Fig. 1c). The differences in RSCU values for each BBSV isolate are indicated in Figure 1d. As shown, most of the amino acids with RSCU > 1.6 were ending with A/U in comparison to G/C ending codons (Fig. 1d).

Table 3. The RSCU value of 59 codons encoding 18 amino acids according to beet soil-borne viruses coat protein genes

Codon	aa	BNYVV	BBSV	BSBV	BVQ	All
UUU	Phe	**1.91**	**1.11**	**1.74**	**1**	**1.44**
UUC		0.09	0.89	0.26	1	0.56
UUA	Leu	**1.96**	**1.24**	**2.4**	0.81	**1.60**
UUG		1.47	0.25	1.07	1.86	1.16
CUU		1.6	1.93	0.51	1.67	1.43
CUC		0	1.57	0.43	0	0.50
CUA		0.65	0.49	0.57	0.52	0.56
CUG		0.32	0.51	1.02	1.14	0.75
AUU	Iie	**1.49**	**1.43**	**1.36**	**1.71**	**1.50**
AUC		0	0.83	0.27	0.43	0.38
AUA		1.51	0.74	1.36	0.86	1.12
GUU	Val	**1.47**	**1.76**	**2.41**	**1.6**	**1.81**
GUC		0.38	0.75	0.57	0	0.43
GUA		0.37	0.36	0.34	0.8	0.47
GUG		1.77	1.13	0.68	1.6	1.30
UCU	Ser	1.34	0.85	0.72	0	0.73
UCC		**1.14**	**1.56**	**2**	**0.72**	**1.36**
UCA		0.9	0.81	0.05	0.36	0.53
UCG		1.2	0.08	0.59	1.81	0.92
AGU		1.41	0.68	1.33	1.81	1.31
AGC		0.01	2.02	1.31	1.29	1.16
CCU	Pro	1.01	1.83	0	0	0.71
CCC		0.5	1.35	0	1.33	0.80
CCA		1.99	0.67	0	1.33	1.00
CCG		**0.5**	**0.15**	**4**	**1.33**	**1.50**
ACU	Thr	**2.08**	**1.37**	**2.26**	**0.7**	**1.60**
ACC		0.95	1.06	0.43	1.08	0.88
ACA		0.72	1.05	0.32	0	0.52
ACG		0.25	0.53	0.99	2.22	1.00

Codon	aa	BNYVV	BBSV	BSBV	BVQ	All
GCU	Ala	**2.22**	**2.41**	**1.53**	**1.45**	**1.90**
GCC		0.37	0.49	1.04	0.36	0.57
GCA		0.92	0.98	1.13	1.25	1.07
GCG		0.5	0.13	0.3	0.94	0.47
UAU	Tyr	**1.54**	**1.02**	**1**	**1.33**	**1.22**
UAC		0.46	0.98	1	0.67	0.78
CAU	His	**1.98**	0	**1.31**	**1.68**	**1.24**
CAC		0.02	0	0.69	0.32	0.26
CAA	Gln	**1.59**	**1.10**	0.79	**1.65**	**1.28**
CAG		0.41	0.90	1.21	0.35	0.72
AAU	Asn	**1.85**	**1.25**	**1.14**	**1.2**	**1.36**
AAC		0.15	0.75	0.86	0.8	0.64
AAA	Lys	**0.97**	**0.74**	**0.57**	**1.8**	**1.02**
AAG		1.03	1.26	1.43	0.2	0.98
GAU	Asp	**1.09**	**1.63**	**1.4**	**1.5**	**1.41**
GAC		0.91	0.37	0.6	0.5	0.60
GAA	Glu	1.21	0.24	1	1.06	0.88
GAG		**0.79**	**1.76**	**1**	**0.94**	**1.12**
UGU	Cys	2	0.72	0	0	0.68
UGC		**0**	**1.27**	**2**	**0**	**0.82**
CGU	Arg	1.78	0.90	0.67	1.12	1.12
CGC		0	1.42	1.09	0.96	0.87
CGA		0.59	0.67	0.97	0.88	0.78
CGG		1.2	0.64	0.8	0.8	0.86
AGA		**1.8**	**0.60**	**1.4**	**1.12**	**1.23**
AGG		0.62	1.76	1.07	1.12	1.14
GGU	Gly	**2.05**	**1.38**	**1.3**	**1.19**	**1.48**
GGC		0.58	0.86	0.5	1.02	0.74
GGA		1.03	1.08	1.44	0.89	1.11
GGG		0.34	0.68	0.76	0.89	0.67

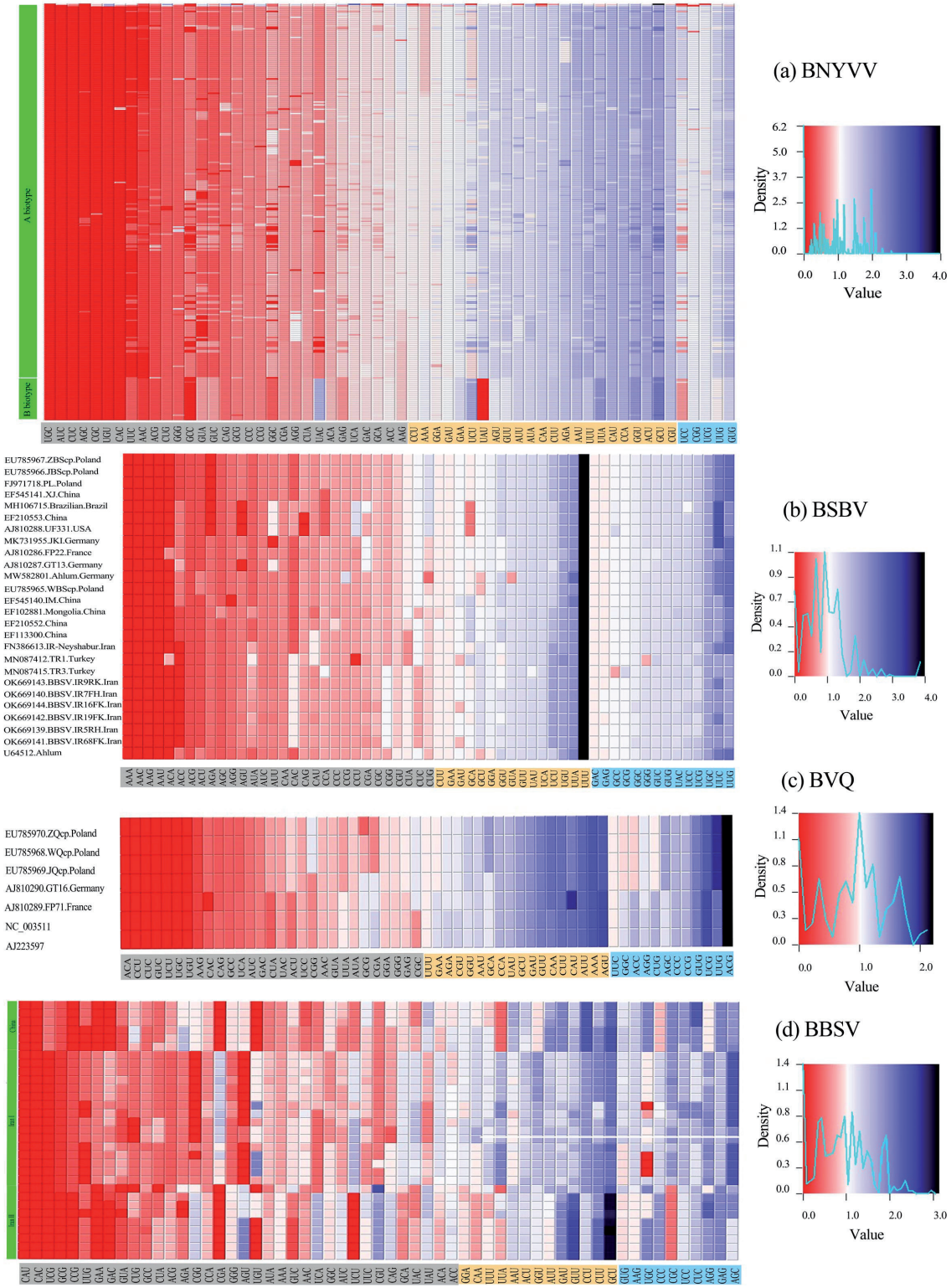

Fig. 1

RSCU value of each codon in the *CP* gene of each beet soil-borne virus isolates is shown (a) BNYVV, (b) BSBV, (c) BVQ, and (d) BBSV Rows indicated the 59 nondegenerate, non-stop codons. Overrepresented codons (RSCU > 1) are shown by blue cells and underrepresented codons (RSCU < 1) by red cells. Codons were grouped into three main clusters by using RSCU values: underrepresented A/U-and G/C ending codons (cluster I, grey bar), A/U ending codons (cluster II, orange bar), and G/C ending codons (cluster III, blue bar).

Codon usage bias of the CP genes

To estimate the degree of codon usage bias within beet soil-borne virus CP sequences, ENC values were calculated. Generally, ENC values differ from 20 (top bias) to 61 (no bias) (He *et al.*, 2019) and the stronger codon usage bias is indicated by a smaller ENC value. Low codon usage bias in all CP coding sequences of beet soil-borne viruses was found with ENC average value of 51.58 ± 1.258; 52.06 ± 3.224; 59.27 ± 0.621; and 60.14 ± 1.249 for BNYVV, BBSV, BVQ, and BSBV, respectively.

Trends in codon usage variation

Based on PCA analysis there is an overlap between strains of the different geographical isolation of BNYVV, indicating a common ancestor (Fig. 2a). Among them, the BNYVV isolates from East Asia were nearest to the origin (Fig. 3). We assume that the origin of the graph coincides with the ancestral (original) virus (Nasrullah *et al.*, 2015; Butt *et al.*, 2016; Chen *et al.*, 2017; Zhang *et al.*, 2019; Konishi 2020). In addition, three Chinese isolates Har4, CBCECP-4 (biotype A), and KYCP-6 (biotype B), were far from other BNYVV isolates (Fig. 2a), which may show the recent emergence of these viruses due to evolutionary founder effects/bottlenecks, or strong selection pressures (for example, host adaptation). The same result was also obtained after analysis of *p25* gene sequences of BNYVV using PCA, which indicates that the BNYVV possibly originated from East Asia (Fig. 4a). Among the East Asians, the isolates from China were older than Japanese isolates (Fig. 4b). Furthermore, among the European isolates the French isolates (MG839239, MG839240), and one Spanish isolate (AY696173) were closer to the origin. This indicated that these isolates possibly are the oldest European BNYVV isolates (Fig. 4c), and confirmed previous studies that showed that BNYVV spread from central and southern Europe to northern and Eastern Europe during a few decades (Asher 1993; Tamada 1999; McGrann *et al.*, 2009). The same result was also inferred for BSBV isolates when the plotted axes based on geographical locations were drawn. This analysis also indicated that the Chinese isolates are closer to the origin than the others (Fig. 2b, Fig. 5). As indicated in Fig. 2c, BBSV isolates are separated into three diverged groups.

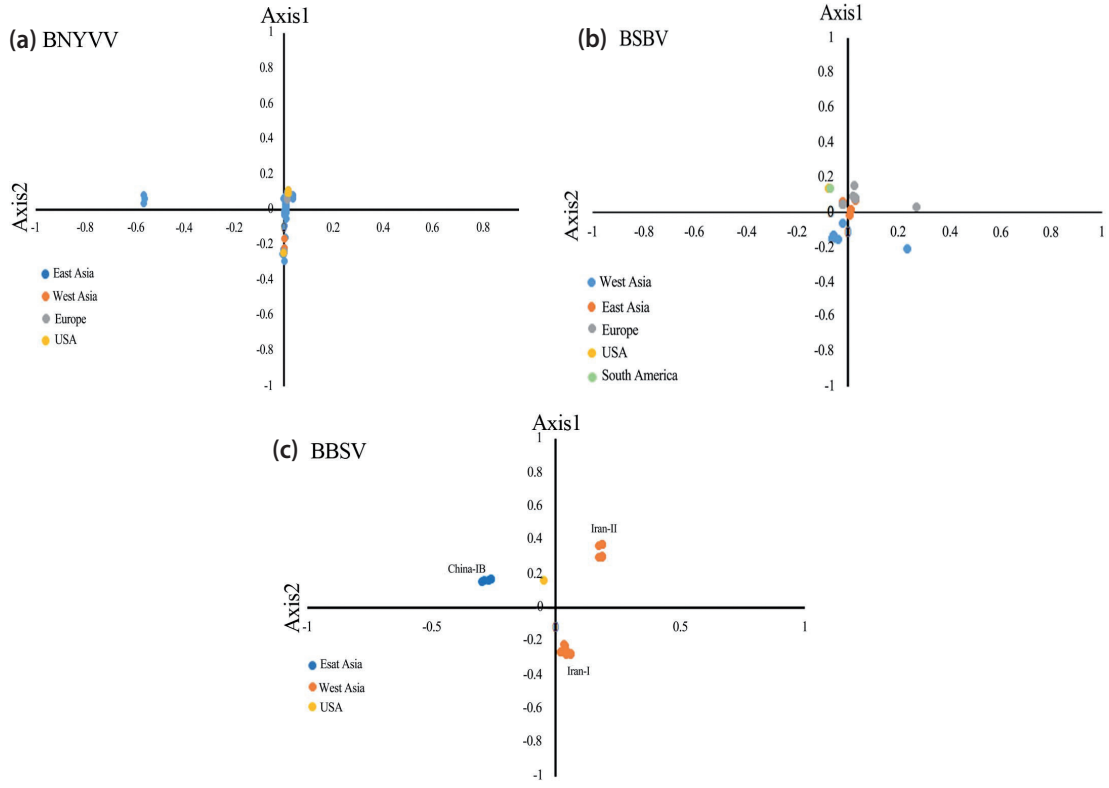

Fig. 2

Principal component analysis (PCA) according to geographical locations in which viruses were isolated
(a) BNYVV; **(b)** BSBV; **(c)** BBSV. East Asia, West Asia, Europe, and the USA isolates are shown in blue, orange, gray, and yellow respectively. For a BSBV isolate from Brazil (South America) the green color is used.

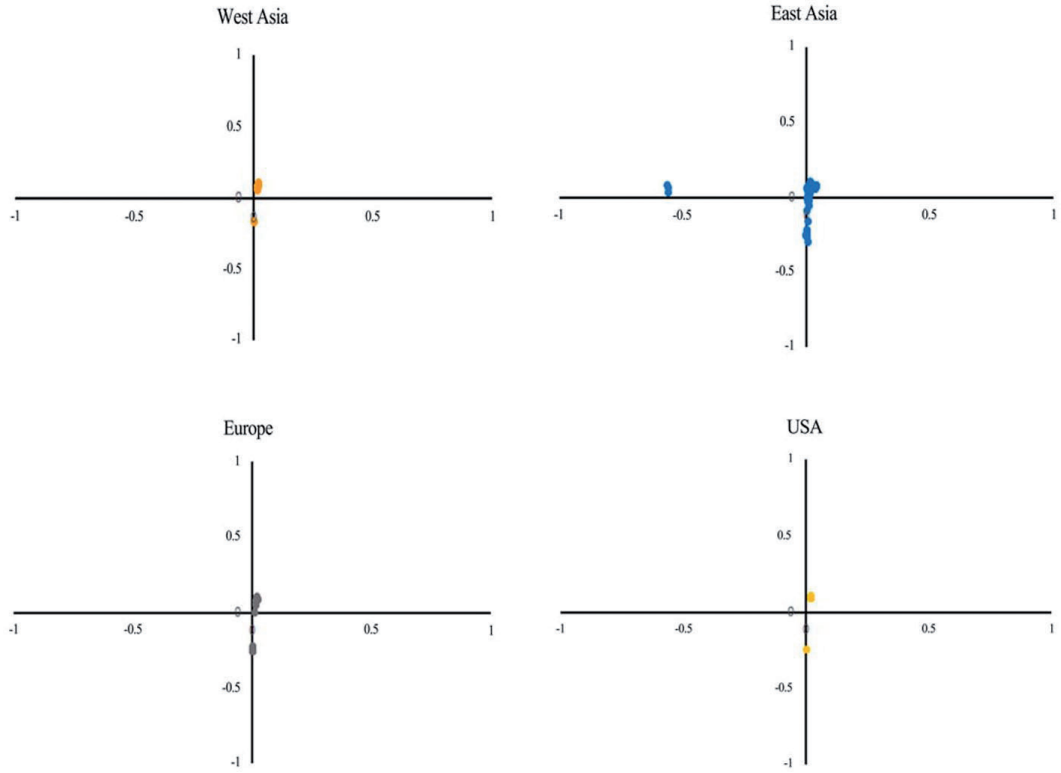

Fig. 3

Principal component analysis (PCA) according to geographical locations in which BNYVV was isolated
East Asia, West Asia, Europe, and USA isolates are shown in blue, orange, gray, and yellow colors respectively.

PR2 biasness analysis

The relation between A3 and U3 content and G3 and C3 content was analyzed by PR2 plots to determine if the biased codon choices were restricted to highly biased *CP* genes of beet soil-borne viruses. There is no bias in the selection or mutation pressure when the plot lies in the center, with both coordinates at 0.5 (Sueoka, 1995). The preference toward codon choices is confirmed by the PR2 plot (Fig. 6). It was indicated that A and U were used more frequently than G and C in the *CP* gene of beet soil-borne viruses. The AU bias in *CP* genes using the PR2 plot demonstrates the preference towards codon choices shaped by natural selection and mutation pressure to variable degrees.

Mutation pressure and natural selection, both play roles in the CUB of beet soil-borne viruses CP genes

ENC-plot using *CP* gene sequences was performed to find out if the patterns of codon usage have been affected by selection pressure or mutation pressure. By ENC values against GC3s values the data show that BNYVV (Fig. 7a), BSBV (Fig. 7b), and BVQ (Fig. 7c) isolates clustered near/on the expected curve, indicating that mutation pressure is more dominant than selection pressure in codon usage patterns of beet soil-borne virus isolates. Using *p25* gene sequences of BNYVV, which were obtained from GenBank data, it was found that selection pressure is more dominant than mutation pressure in the coding usage bias of this gene (Fig. 8). In addition, BBSV isolates from different geographical isolations fall slightly below on the left side of the expected curve, indicating that the natural selection pressure was more important than mutation pressure in BBSV isolates (Fig. 7d).

Codon usage host adaptation

The CAI analysis was done for assessment of the codon usage optimization and host adaptation of beet soil-borne viruses. The average CAI values of the CP coding sequences were 0.778, 0.777, and 0.734 for BBSV, BNYVV, and BVQ, respectively. To understand how the host's codon usage patterns affect the viral codon usage patterns, SiD analysis was performed. The SiD value of BBSV (0.171) was higher than BNYVV (0.123), BSBV (0.117),

and BVQ (0.099) indicating that during beet soil-borne virus evolution, sugar beet has a greater impact on the BBSV. This result is in agreement with neutrality and ENC analysis, which indicates that selection pressure on BBSV was greater than on the BNYVV, BSBV, and BVQ. In addition, similarity index analysis indicated the common vector *P. betae* has a profound effect on shaping RSCU patterns of BNYVV (0.209), followed by BSBV (0.158), and BVQ (0.134).

Discussion

The objectives of this study were to better understand the sequence diversity and genetic structure of soil-borne viruses infecting sugar beet using different approaches. In this respect, new Iranian CP sequences of BNYVV, BSBV, and BBSV isolates were sequenced. Soil samples from West Iran were collected because the sugar beet fields in these areas suffer mostly from beet yield reduction caused by soil-borne viruses as indicated by ELISA tests. The highest incidence was found for BNYVV (72.4%), followed by BSBV (38.0%), and BBSV (19.0%). No BVQ positive sample were detected in this study. BNYVV has the highest impact, but it can mostly be controlled by genetic resistance. The effects of BSBV and BVQ are under evaluation, but it is speculated that they play a minor role and so far, genetic resistances as a control measure are not known (Biancardi and Tamada, 2016). Phylogenetic analyses clustered the new Iranian BNYVV isolates into main group A genotype as it is dominant in the rest of the world. The higher fitness or transmission efficiency by vector may be related to the wide distribution and prevalence of genotype A than genotype B. The overall low variability observed in BNYVV and two pomovirus (BSBV and BVQ) isolates worldwide suggests high genetic stability. This could be explained, at least in part, by the ecological bottleneck in which host plants and vectors are continuously maintained. Genetic bottlenecks are evolutionary events that reduce the genetic variation of a population and may occur at different

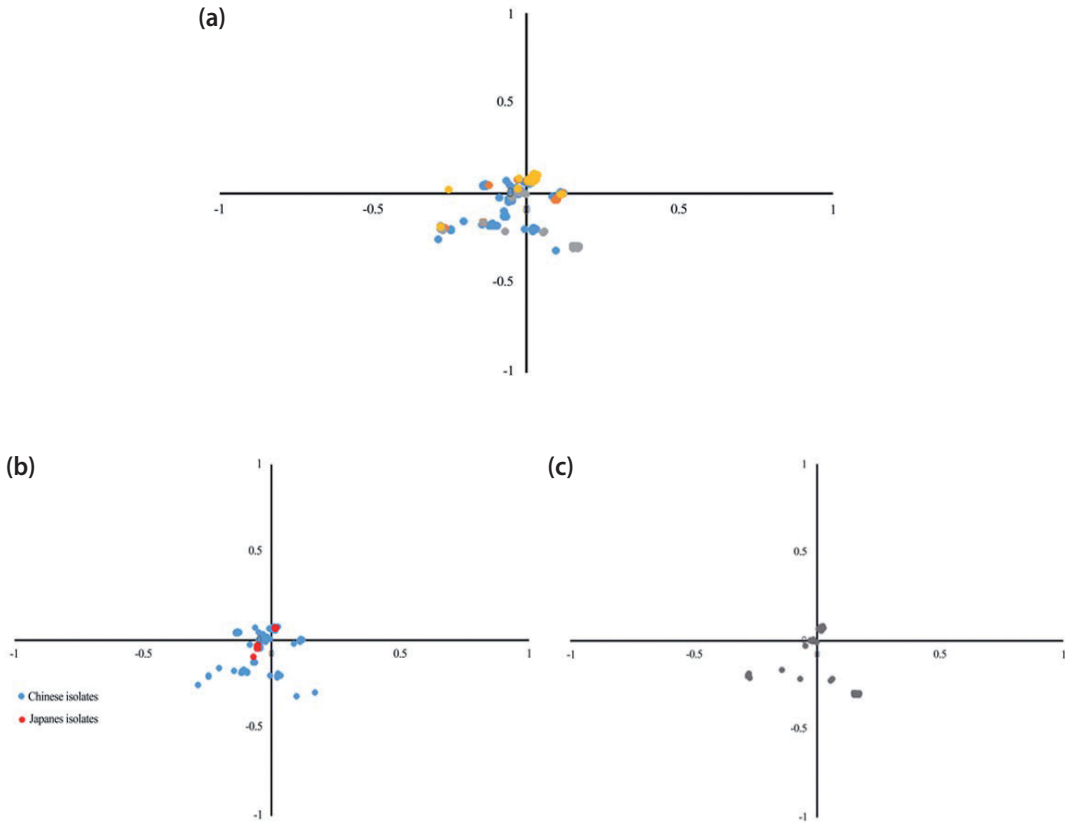

Fig. 4

Principal component analysis (PCA) of BNYVV

(a) The principal component analysis (PCA) of the BNYVV *p25* gene sequences according to geographical locations; **(b)** PCA analysis according to East Asia isolates (in this plot Japanese isolates indicated by red color) **(c)** PCA analysis according to Europe isolates. East Asia, West Asia, Europe, and South American isolates are shown in blue, orange, gray, and yellow colors respectively.

points during the life cycle of plant RNA viruses. Genetic bottlenecks experimentally have been demonstrated in plant virus populations during systemic movement within the plant and horizontal transmission from plant to plant by vectors (Ali *et al.*, 2006). Similarly, one may expect that the transmission of BSBV by *P. betae* can impose a bottleneck and reduce the genetic variation of the virus population. The separation of Iranian BBSV isolates into three groups indicates that numerous independent infections have happened, likely from isolated sources from the unknown host(s) related together through the common viral vector *Olpidium*. Since the viruliferous resting spores of the vector in soils are the main factor to transfer and spread the beet soil-borne viruses (Biancardi and Tamada 2016), the transmission of BBSV from unidentified hosts to sugar beet fields may be considered as an important factor for its evolution.

Identification of codon usage patterns provides important information about the host-pathogen co-evolution, such as adaptation of pathogens to hosts and molecular evolution of genes (He *et al.*, 2017). In comparison with eukaryotic and prokaryotic organisms, the importance of codon usage bias in the evolution of plant viruses is less considered. It has previously been indicated that codon usage bias, or preference for one type of codon over another, can be significantly influenced by overall

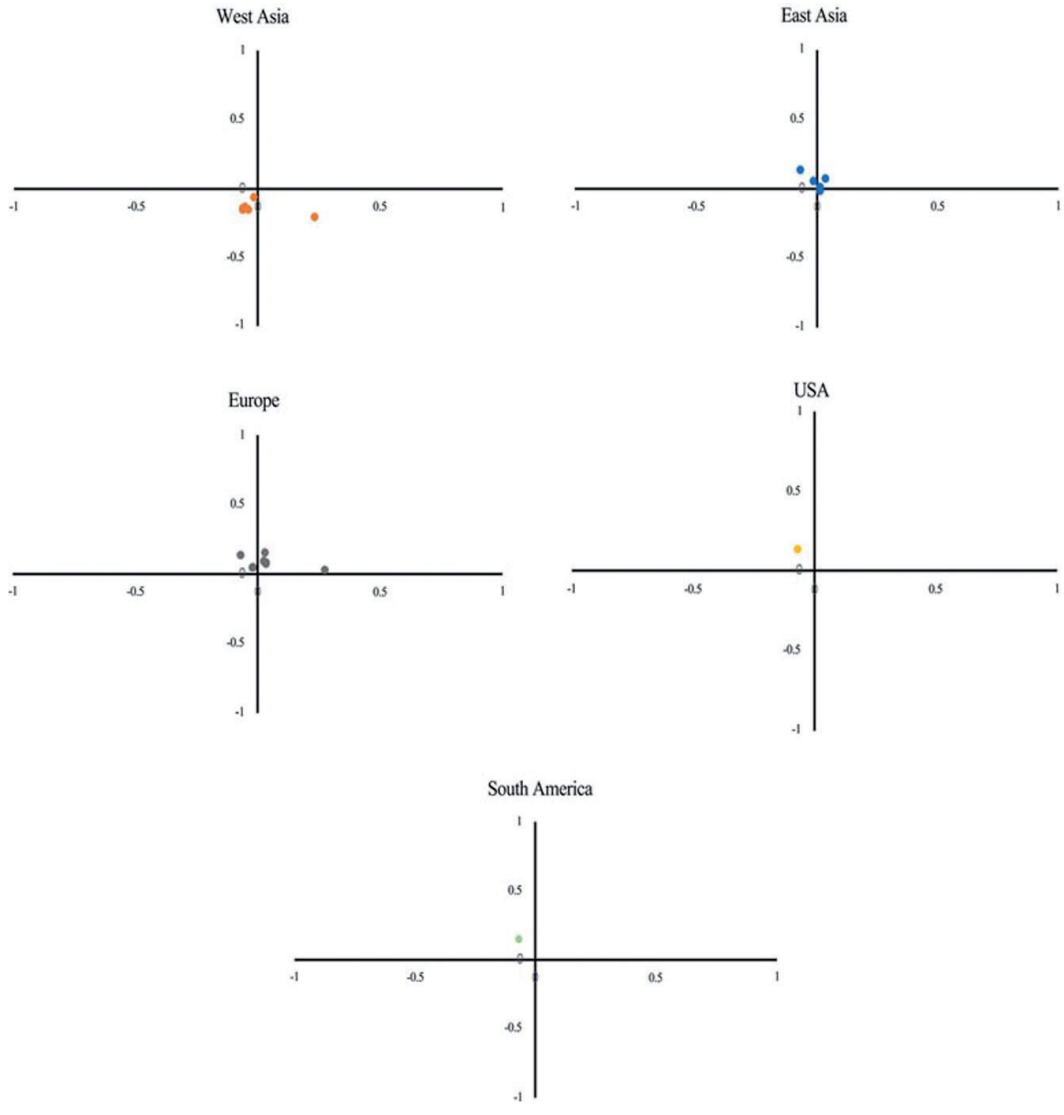

Fig. 5

Principal component analysis (PCA) according to geographical locations in which BSBV was isolated
East Asia, West Asia, Europe, and South American isolates are shown in blue, orange, gray, and yellow colors respectively.

genomic composition (Gu *et al.*, 2020). The significant role of nucleotide compositional constraints in shaping the codon usage patterns in different DNA and RNA virus genomes have been reported (Sharp and Li 1986a,b; Xu *et al.*, 2008; Wong *et al.*, 2010; Liu *et al.*, 2012; He *et al.*, 2017, 2019; Gu *et al.*, 2020). RSCU and Nucleotide composition analyses indicated that selection of the preferred codons has been influenced by composition of A and U, which considers the existence of mutation pressure. In beet soil-borne CP sequences, RSCU analyses indicated that the preferred codons have been mostly influenced by compositional constraints (U and A) and within *CP* genes, codon preferences differ (Fig. 1a-1d). The RSCU values of less than 1 were most frequently seen in CGN/NCG codons, showing a strong CpG deficiency or suppression (Fig. 1a-1d), which is indicative of selection pressure acting on beet soil-borne viruses in shaping codon usage. Lower CpG sites are generally preserved in RNA viruses to avoid stimulation of innate immune responses and also to emulate the host's codon usage as an optimization to the available tRNAs pool (Kumar *et al.*, 2018). The other dinucleotide commonly underrepresented in RNA virus genomes is UpA (i.e. UUA, CUA, AUA, GUA, UAU, and UAC) (Bera *et al.*, 2017). However, for *CP* genes of beet soil-borne viruses, UpA frequency was not underrepresented, and the reason is attributed to A nucleotide-rich *CP* gene. Unbiased use of UpA indicates that selection pressure

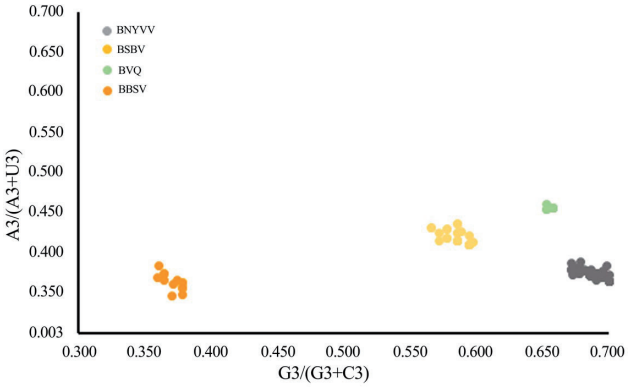

Fig. 6

The AT [A3%/(A3% + T3%)] and GC [G3%/(G3% + C3%)] bias of the beet soil-borne viruses *CP* gene sequences

The BNYVV, BSBV, BVQ, and BBSV are shown by blue, red, and green dots, respectively.

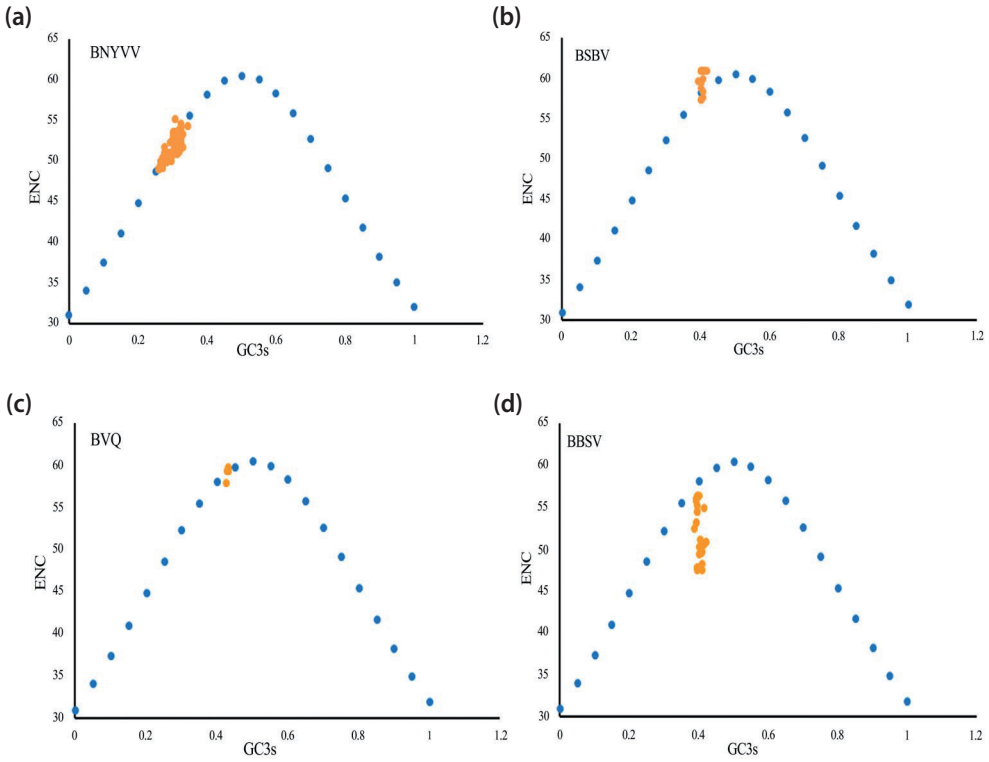

Fig. 7

ENC-plot analysis of the beet soil-borne *CP* gene sequences with ENC curve drawn against GC3s

The standard curve was plotted using the codon usage bias (calculated by the GC3s composition only) indicated by blue points. **(a)** BNYVV, **(b)** BSBV, **(c)** BVQ, **(d)** BBSV.

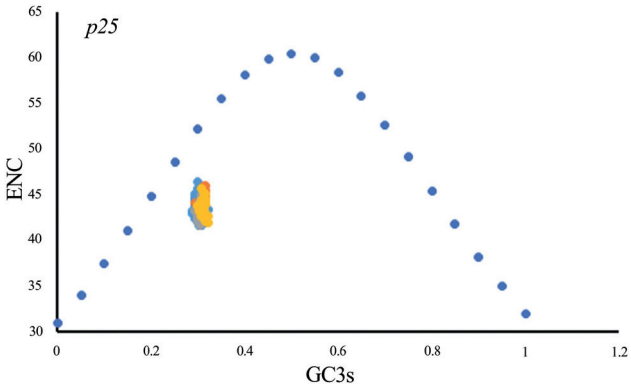

Fig. 8

ENC-plot analysis of the BNYVV *p25* gene sequences with ENC curve drawn against GC3s, according to geographical isolation
A standard curve was plotted using the codon usage bias (calculated by the GC3s composition only) indicated by blue points. East Asia, West Asia, Europe, and American isolates are shown in blue, orange, gray, and yellow colors, respectively.

leading to low UpA frequencies is not actively involved in the codon usage patterns of these viruses, rather, these patterns are mainly governed by compositional constraint (Kumar *et al.*, 2018).

The low ENC values represent an approximately constant and conserved CP composition. Natural selection may be induced by weak or low codon usage bias when the viruses want to adapt to the hosts (Barrett *et al.*, 2006; Shi *et al.*, 2013; Zhang *et al.*, 2013). As inferred for beet soil-borne viruses in this study, the lower codon usage bias has been previously reported for some plant viruses including begomoviruses (Xu *et al.*, 2008), Rice stripe virus (He *et al.*, 2017), Potato virus M (He *et al.*, 2019), Papaya ringspot virus (Chakraborty *et al.*, 2015), Citrus tristeza virus (Biswas *et al.*, 2019). Among a virus population with an RNA genome, subpopulations with a faster replication have more chance to increase because both the virus and host use a common resource for their translational procedures. However, the RNA-dependent RNA polymerase has no 3'–5' proof-reading property, and the virus variants with a higher rate of replication sometimes resulted in the introduction of detrimental mutations and decreasing in population fitness. A lower replication rate results in higher accuracy and fidelity and subsequently increases the fitness of the virus population. Therefore, low codon usage bias of RNA viruses has a benefit for efficient replication in the host by decreasing the competition between the host and virus in utilizing the synthesis systems (Jenkins and Holmes, 2003).

To find out the trends in codon usage variation among coding CP sequences of beet soil-borne viruses, we plotted principal axes according to geographical isolation for each virus. Clustering of the majority of BNVYY and BSBV isolates from East Asia near to origin may illustrate the old dispersal area of these viruses (Fig. 2). PCA analysis confirmed the previous proposal that BNYVV originally evolved in East Asia and has recently become a pathogen of cultivated sugar beet (Chiba *et al.*, 2011). The same results can be inferred for BSBV (Fig. 2b) indicating that this pomovirus might diverge from a common ancestor in some regions of East Asia. BBSV isolates separated into three groups and perhaps independently evolved in three clusters after distribution on principal axes (Fig. 2c). The grouping of diverse BBSV lineages that are separated by thousands of miles within a single main group (Iran-I, Chines, and USA isolates), as well as grouping of closely related isolates into distinct clusters (Iran-I and Iran-II), indicates an important role of the flexibility and movability of BBSV's natural hosts (e.g. sugar beet and chytrid vector). There was not enough data for PCA analysis about BVQ.

Normally, the balance of selection and mutation pressures plays an important role in the codon usage patterns in eukaryotes and prokaryotes (Wu *et al.*, 2007, 2015; Wang *et al.*, 2016). The results of the PR2 and ENC plots indicated that codon usage of beet soil-borne viruses is influenced by natural selection and mutation pressure to variable degrees. CAI and SiD analysis reflected the interplay of codon usage between beet soil-borne viruses, their host, and vector, which influence viral fitness, survival, and evolution. Compared with sugar beet, *P. betae* has a greater effect on shaping RSCU patterns, as concluded from the similarity index analysis. As *P. betae* and *O. brassica* are considered as the natural reservoir and host for beet soil-borne viruses, it makes sense that the virus has evolved its genomic properties to a steady level to better adapt to its primary host's condition.

This study showed that overall codon usage patterns within *CP* genes of sugar beet soil-borne viruses are slightly biased. A low CUB of RNA viruses has an advantage for efficient replication in the host cells by reducing the competition between the virus and host in using the synthesis machinery. The evolution of beet soil-borne viruses perhaps reflects a dynamic process of mutation and natural selection to adapt their codon usage to different environments, hosts, and vectors. Since the viruliferous resting vector spores are transferred via the soils, therefore the spread of them from unrecognized natural hosts to sugar beet fields may have been a major factor in their evolution. This research is the first study of codon usage analysis of beet soil-borne viruses and increased our knowledge about the mechanisms that support codon usage and their evolution.

References

Adams MJ, Adkins S, Bragard C, Gilmer D, Li D, MacFarlane SA, Wong S-M, Melcher U, Ratti C, Ryu KH (2017): ICTV Virus taxonomy profile: Virgaviridae. J. Gen. Virol. 98, 1999–2000. https://doi.org/10.1099/jgv.0.000884

Ali A, Schneider WL, Sherman DJ, Gray S, Smith D, Roosinck MJ (2006): Analysis of genetic bottlenecks during horizontal transmission of cucumber mosaic virus. J. Virol. 80, 8345–8350. https://doi.org/10.1128/JVI.00568-06

Asher MJC (1993): Rhizomania. In: Cooke DA, Scott RK (Eds.) The sugar beet crop: science into practice. Chapman & Hall, London, pp. 311–346. https://doi.org/10.1007/978-94-009-0373-9_9

Barrett JW, Sun Y, Nazarian SH, Belsito TA, Brunetti CR, McFadden G (2006): Optimization of codon usage of poxvirus genes allows for improved transient expression in mammalian cells. Virus Genes 33, 15–26. https://doi.org/10.1007/s11262-005-0035-7

Bera BC, Virmani N, Kumar N, Anand T, Pavulraj S, Rash A, Elton D, Rash N, Bhatia S, Sood R, Singh RK, Tripathi BN (2017): Genetic and codon usage bias analyses of polymerase genes of equine influenza virus and its relation to evolution. BMC Genom. 18, 652 https://doi.org/10.1186/s12864-017-4063-1

Biancardi E, Tamada T (2016): Rhizomania. Springer International Publishing. 281 pp. https://doi.org/10.1007/978-3-319-30678-0

Biswas K, Palchoudhury S, Chakraborty P, Bhattacharyya U, Ghosh D, Debnath P, Ramadugu C, Keremane ML, Khetarpal RK, Lee RF (2019): Codon usage bias analysis of Citrus tristeza virus: Higher codon adaptation to citrus reticulata host. Viruses 11, 331. Epub 2019 April 8. https://doi.org/10.3390/v11040331

Borodynko N, Rymelska N, Hasiów-Jaroszewska B, Pospieszny H (2009): Molecular characterization of three soil-borne sugar beet-infecting viruses based on the coat protein gene. J. Plant Path. 91, 191–193.

Butt AM, Nasrullah I, Qamar R, Tong Y (2016): Evolution of codon usage in Zika virus genomes is host and vector specific. Eme. Microb. Infect. 5, e107. Epub 2016 October 12. https://doi.org/10.1038/emi.2016.106

Callaway A, Giesman-Cookmeyer D, Gillock ET, Sit TL, Lommel SA (2001): The multifunctional capsid proteins of plant RNA viruses. Annu. Rev. Phytopathol. 39, 419–460. https://doi.org/10.1146/annurev.phyto.39.1.419

Canova A (1959): Appunti di pathologia della barbabietola. Inform. Fito. 9, 390–396.

Chakraborty P, Das S, Saha B, Sarkar P, Karmakar A, Saha A, Saha D, Saha A (2015): Phylogeny and synonymous codon usage pattern of Papaya ringspot virus coat protein gene in the sub-Himalayan region of north-east India. Canadian J. Microbiol. 61, 555–564. https://doi.org/10.1139/cjm-2015-0172

Chen Y, Li X, Chi X, Wang S, Ma Y, Chen J (2017): Comprehensive analysis of the codon usage patterns in the envelope glycoprotein E2 gene of the classical swine fever virus. PLoS ONE 12, e0183646. Epub 2017 September 7. https://doi.org/10.1371/journal.pone.0183646

Chiba S, Kondo H, Miyanishi M, Andika IB, Han C, Tamada T (2011): The evolutionary history of beet necrotic yellow vein virus deduced from genetic variation, geographical origin and spread, and the breaking of host resistance. Mol. Plant-Micro. Inter. 24, 207–218. Epub 2011 January 7. https://doi.org/10.1094/MPMI-10-10-0241

Clark MF, Adams AN (1977): Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. J. Gen. Virol. 34,475–483. Epub 1977 March 1. https://doi.org/10.1099/0022–1317–34–3–475 https://doi.org/10.1099/0022-1317-34-3-475

Comeron JM, Aguade M (1998): An evaluation of measures of synonymous codon usage bias. J. Mol. Evol. 47, 268–274. https://doi.org/10.1007/PL00006384

Cui X (1988): An icosahedral virus found in sugar beet. J. Xinjiang Shihezi Agri. Coll. 10, 73–78.

Davino S, Panno S, Arrigo M, La Rocca M, Caruso AG, Lo Bosco G (2017): Planthology: an application system for plant diseases management. Chem. Engi. Trans. 58, 619–624.

Duffus JE, Whitney ED, Larsen RC, Liu HY, Lewellen RT (1984): First report in Western hemisphere of rhizomania of sugar beet caused by beet necrotic yellow vein virus. Plant Dis. 68, 251.

Elena S, Fraile A, Garcıʻa-Arenal F (2014): Evolution and emergence of plant viruses. Adva. Virus Res. 88, 161–191. https://doi.org/10.1016/B978-0-12-800098-4.00003-9

Farzadfar S, Pourrahim R (2019): Positive selection and recombination shaped the large genetic differentiation of Beet black scorch virus population. PLoS ONE 14, e0215574. Epub 2019 April 25. https://doi.org/10.1371/journal.pone.0215574

Gao J, Deng F, Zhai H, Liang X, Liu Y (1983): The occurrence of sugar beet rhizomania caused by beet necrotic yellow vein virus in China. Acta Phytopathol. Sinica 13, 1–4.

Garcia-Arenal F, Fraile A, Malpica JM (2001): Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39, 157–186. https://doi.org/10.1146/annurev.phyto.39.1.157

Gu H, Chu DKW, Peiris M, Poon LLM (2020): Multivariate analyses of codon usage of SARS-CoV-2 and other betacoronaviruses. Virus Evol. 6. https://doi.org/10.1093/ve/veaa032

Hall TA (1999): BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hamdorf G, Lesemann DE, Weidemann HL (1977): Untersuchungen über die Rizomania-Krankheit an Zuckerrüben in der Bundesrepublik Deutschland. J. Phytopath. 90, 9–103. https://doi.org/10.1111/j.1439-0434.1977.tb03223.x

He M, Guan SY, He Ch-Qi (2017): Evolution of rice stripe virus. Mol. Phylo. Evol. 109, 343–350. https://doi.org/10.1016/j.ympev.2017.02.002

He Z, Gan H, Liang X (2019): Analysis of synonymous codon usage bias in Potato virus M and its adaption to hosts. Viruses 11, E752. https://doi.org/10.3390/v11080752

Henry C (1996): Rhizomania-its effect on sugar beet yield in the UK. British Sugar Beet Rev. 64, 24–26.

Hill SA, Torrance L (1989): Rhizomania disease of sugar beet in England. Plant Path. 38, 114–122. https://doi.org/10.1111/j.1365-3059.1989.tb01437.x

Izadpanah K, Hashemi P, Kamran R, Pakniat M, Sahandpour A, Masumi M (1996): Widespread occurrence of rhizomania-like disease of sugar beet in Fars. Iranian J. Plant Path. 32, 155.

Jenkins GM, Holmes EC (2003): The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 92, 1–7. https://doi.org/10.1016/S0168-1702(02)00309-X

Koenig R, Valizadeh J (2008): Molecular and serological characterization of an Iranian isolate of Beet black scorch virus. Arch. Virol. 153, 1397–1400. https://doi.org/10.1007/s00705-008-0121-y

Konishi T (2020): Principal component analysis of coronaviruses reveals their diversity and seasonal and pandemic potential. PLoS ONE 15, e0242954. Epub 2020 December 3. https://doi.org/10.1371/journal.pone.0242954

Kouyeas H (1979): The rhizomania of sugar beet. Anna. Inst. Phytopath. Benaki, Greece 12, 151–153.

Kumar N, Kulkarni DD, Lee B, Kaushik R, Bhatia S, Sood R, Pateriya AK, Bhat S, Singh VP (2018): Evolution of codon usage bias in Henipaviruses is governed by natural selection and is host-specific. Viruses 604. https://doi.org/10.3390/v10110604

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018): MEGAX: Molecular evolutionary genetics analysis across Computing Platforms Sudhir. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Liu XS, Zhang YG, Fang YZ, Wang Y-Lu (2012): Patterns and influencing factor of synonymous codon usage in porcine circovirus. Virol. J. 9, 1–9. https://doi.10.1186/1743–422X-9–68 https://doi.org/10.1186/1743-422X-9-68

McGrann GRD, Grimmer MK, Mutasa-Göttgens ES, Stevens M (2009): Progress towards the understanding and control of sugar beet rhizomania disease. Mol. Plant Pathol. 10, 129–141. https://doi.org/10.1111/j.1364-3703.2008.00514.x

Meunier A, Schmit JF, Stas A, Kutluk N, Bragard C (2003): Multiplex reverse transcription-PCR for simultaneous detection of Beet necrotic yellow vein virus, Beet soil borne virus, and Beet virus Q and their vector Polymyxa betae Keskin on sugar beet. Appl. Environ. Microbiol. 69, 2356–2360. https://doi.org/10.1128/AEM.69.4.2356-2360.2003

Nasrullah I, Butt AM, Tahir S, Idrees M, Tong Y (2015): Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on Marburg virus evolution. BMC Evol. Biol. 15, 174. https://doi.org/10.1186/s12862-015-0456-4

Puigbo P, Bravo IG, Garcia-Vallve S (2008): CAIcal: a combined set of tools to assess codon usage adaptation. Biol. Dir. 3, 38. https://doi.org/10.1186/1745-6150-3-38

Putz C, Vuittenez A (1974): Observation des particules virales chez des betteraves presentant, en Alsace, des symptomes de ‚Rhizomanie'. Anna. de Phytopath. 6, 129–138.

Sharp PM, Li WH (1986a): Codon usage in regulatory genes in Escherichia coli does not reflect selection for ‚rare' codons. Nuc. Acids Res. 14, 7737–7749. https://doi.org/10.1093/nar/14.19.7737

Sharp PM, Li WH (1986b): An evolutionary perspective on synonymous codon usage in unicellular organisms. J. Mol. Evol. 24, 28–38. https://doi.org/10.1007/BF02099948

Sharp PM, Tuohy TMF, Mosurski KR (1986): Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. Nuc. Acids Res. 14, 5125–5143. https://doi.org/10.1093/nar/14.13.5125

Shi SL, Jiang YR, Liu YQ, Xia RX, Qin L (2013): Selective pressure dominates the synonymous codon usage in parvoviridae. Virus Genes 46, 10–19. https://doi.org/10.1007/s11262-012-0818-6

Suárez MB, Grondona I, García-Benavides P, Monte E, García-Acha I (1999): Characterization of beet necrotic yellow vein furovirus from Spanish sugar beets. Interac. Microbiol. 2, 87–92.

Sueoka N (1995): Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. J. Mol. Evol. 40, 318–325. https://doi.org/10.1007/BF00163236

Sueoka N (1999): Translation-coupled violation of Parity Rule 2 in human genes is not the cause of heterogeneity of the DNA G+C content of third codon position. Gene 238, 53–58. https://doi.org/10.1016/S0378-1119(99)00320-0

Tamada T (1999): Benyvirus. In: Granoff A, Webster R (Eds.) Encyclopedia of virology, 2nd edn. Academic, London UK, pp. 154–160. https://doi.org/10.1006/rwvi.1999.0325

Tamada T (2002): Beet necrotic yellow vein virus. CMI/AAB descriptions of plant viruses No. 391, Wellesbourne, UK.

Tamada T, Baba T (1973): Beet necrotic yellow vein virus from rhizomania affected sugar beet in Japan. Annu. Phytopath. Soc. Japan 39, 325–332. https://doi.org/10.3186/jjphytopath.39.325

Wang H, Liu S, Zhang B, Wei W (2016): Analysis of synonymous codon usage bias of zika virus and its adaption to the hosts. PLoS ONE 11, e0166260. Epub 2016 November 28. https://doi.org/10.1371/journal.pone.0166260

Weiland JJ, Van Winkle D, Edwards MC, Larson RL, Shelver WL, Freeman TP, Liu HY (2007): Characterization of a U.S. isolate of Beet black scorch virus. Phytopath. 97, 1245–1254. https://doi.org/10.1094/PHYTO-97-10-1245

Whitney ED, Duffus E (Eds.) (1998): Compendium of beet diseases and insects, 3rd ed. American Phytopathological Society, St. Paul, Minnesota, USA.

Wong EH, Smith DK, Rabadan R, Peiris M, Poon LL (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, 23–29. https://doi.org/10.1016/0378-1119(90)90491-9

Wu XM, Wu SF, Ren DM, Zhu YP, He FC (2007): The analysis method and progress in the study of codon bias. Hereditas (Beijing) 29, 420–426. https://doi.org/10.1360/yc-007-0420

Wu Y, Zhao D, Tao J (2015): Analysis of codon usage patterns in Herbaceous Peony (Paeonia lactiflora Pall.) based on

transcriptome data. Genes 6, 1125–1139. https://doi.org/10.3390/genes6041125

Xu X-Zh, Liu Qi-po, Fan L-ji, Cul Xi-fe, Zhou Xu-pi (2008): Analysis of synonymous codon usage and evolution of begomoviruses. J. Zhejiang Univ. Sci. 9, 667–674. https://doi.org/10.1631/jzus.B0820005

Zhang KW, Wang SL, Cai ZN, Zhang HS (1996): Preliminary study on beet black scorch disease in Ningxia province. China Sugar Beet 2, 3–8.

Zhang Z, Dai W, Wang Y, Lu C, Fan H (2013): Analysis of synonymous codon usage patterns in torque teno sus virus 1 (TTSuV1). Arch. Virol. 158, 145–154. Epub 2012 September 26. https://doi.org/10.1007/s00705-012-1480-y

Zhang W, Zhang L, He W, Zhang X, Wen B, Wang C, Xu Q, Li G, Zhou J, Veit M, Su S (2019) Genetic evolution and molecular selection of the HE gene of Influenza C virus. Viruses 11, 167. https://doi.org/10.3390/v11020167

Zhou JH, Zhang J, Sun DJ, Ma Q, Chen H-Tai, Ma Li-na, Ding Y-zh, Liu Y-sh (2013): The distribution of synonymous codon choice in the translation initiation region of dengue virus. PLoS ONE 8, e77239. Epub 2013 October 25. https://doi.org/10.1371/journal.pone.0077239