


Evolutionary Rates and Phylogeographical Analysis of *Odontoglossum* Ringspot Virus Based on the 166 Coat Protein Gene Sequences

Zhen He , Tingting Dong, Weiwen Wu, Wen Chen, Xian Liu, and Liangjun Li*

School of Horticulture and Plant Protection, Yangzhou University, Yangzhou 225009, China

(Received on April 22, 2019; Revised on July 10, 2019; Accepted on August 19, 2019)

Odontoglossum ringspot virus (ORSV) is a member of the genus *Tobamovirus*. It is one of the most prevalent viruses infecting orchids worldwide. Earlier studies reported the genetic variability of ORSV isolates from Korea and China. However, the evolutionary rate, timescale, and phylogeographical analyses of ORSV were unclear. Twenty-one coat protein (CP) gene sequences of ORSV were determined in this study, and used them together with 145 CP sequences obtained from GenBank to infer the genetic diversities, evolutionary rate, timescale and migration of ORSV populations. Evolutionary rate of ORSV populations was 1.25×10^{-3} nucleotides/site/y. The most recent common ancestors came from 30 year ago (95% confidence intervals, 26-40). Based on CP gene, ORSV migrated from mainland China and South Korea to Taiwan island, Germany, Australia, Singapore, and Indonesia, and it also circulated within east Asia. Our study is the first attempt to evaluate the evolutionary rates, timescales and migration dynamics of ORSV.

Keywords : evolutionary rates, gene flow, migration, *Odontoglossum ringspot virus*

Handling Editor : Yoon, Ju-Yeon

*Corresponding author.

Phone) +86-514-87979394, FAX) +86-514-87347537

E-mail) ljli@yzu.edu.cn

ORCID

Zhen He

<https://orcid.org/0000-0003-2460-1583>

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Articles can be freely viewed online at www.ppjonline.org.

Due to the high error rates of RNA-dependent RNA polymerases (RdRp), RNA viruses harbor abundant genetic variability in their natural populations (Domingo and Holland, 1997; García-Arenal et al., 2001). The molecular evolution of plant viruses will help us to understand the phylogenetic relationships, epidemiological routes, population structure, the underlying evolutionary mechanisms of RNA viruses, and their ‘emergence’ as new epidemics, simultaneously, this understanding can be essential for designing strategies for controlling viruses (García-Arenal et al., 2001; Gibbs and Ohshima, 2010).

Odontoglossum ringspot virus (ORSV) belonged to the genus *Tobamovirus*, which is one of the most prevalent and economically important viruses infecting orchids worldwide (Jensen and Gold, 1951; Wong et al., 1994; Zettler et al., 1990). It has rod-shaped particles with a positive-sense single-stranded RNA genome. Similar to other tobamoviruses, ORSV encodes at least four proteins, such as the RdRp (126 kDa/183 kDa), movement protein (MP; 33 kDa MP), and coat protein (CP, 18 kDa) (Chng et al., 1996; Ryu and Park, 1995). ORSV induces ringspots, diamond-shaped mottle, and streaks or striped mosaicism symptom on leaves; and crinkle, color breaking and ringspots on orchids flowers (Jensen and Gold, 1951; Wong et al., 1994; Zheng et al., 2010). ORSV can be transmitted by mechanical inoculation with young infected leaves sap.

Earlier studies reported the genetic variability of ORSV isolates from Korea (Ajikuttira et al., 2002; Yoon et al., 2011) and China (Rao et al., 2015). These showed high sequence conservation in CP gene. However, the evolutionary rate, timescale, and geographical adaptation analyses of ORSV were unclear. We report here that we have determined 21 CP sequences of ORSV collected in China and used them together with 145 CP sequences downloaded from GenBank to infer the evolutionary timescale, the degree of divergence and migration of

ORSV populations in different countries.

Materials and Methods

Virus isolates. *Cymbidium goeringii* cultivar ‘Runmei’ and ‘Wang zi’ showing diamond-shaped mottle symptom were collected from commercial orchid plantations in Jiangsu province of China during 2017. The fresh leaves were assayed by reverse transcription polymerase chain reaction (RT-PCR) and double antibody sandwich enzyme-linked immunosorbent assay (Qiaodu, Shanghai, Japan), and the infected leaves were stored at -80°C until use. Details of the Chinese ORSV isolates, their place of origin, year of collection, original variety, symptom, and accession codes in DDBJ/EMBL/GenBank databases are shown in Supplementary Table 1, together with details of the isolates used in the analyses and for which CP sequences have already been reported.

Viral RNA and sequencing. We determined 21 CP sequences of ORSV collected in Jiangsu province of China. The viral RNAs were extracted from ORSV-infected *C. goeringii* leaves using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). For cDNA synthesis and amplification of the CP gene, we used a PrimeScript II High Fidelity One Step RT-PCR kit (Takara, Dalian, China). The previously reported primers described by Rao et al. (2015) were used to amplify CP gene. The RT-PCR products were separated by electrophoresis in agarose gels and purified using the QIAquick Gel Extraction Kit (Qiagen, Beijing, China). The resulting fragments were cloned into the pGEM-T vector (Promega, Madison, WI, USA). The recombinant DNA was transformed into *Escherichia coli* DH5 α . Sequences of CP gene of each isolate were determined using at least three independent cloned plasmids. Each clone was sequenced by primer walking in both directions using an automated DNA sequencer (ABI PRISM™ 3730XL DNA Analyzer). Sequence data were assembled using BioEdit version 5.0.9 (Hall, 1999).

Recombination analyses. The CP gene sequence of *Bell pepper mottle virus* (BPMoV, accession code: DQ355023) (Rhie et al., 2007) was used to align the ORSV CP sequences as BLAST searches had shown it to be the sequences in the international sequence databases most closely and consistently related to ORSV. Those were aligned with the CP sequences of BPMoV as the outgroup using CLUSTAL X2 (Larkin et al., 2007). However, this procedure resulted in some gaps that were not in multiples

of 3 nucleotides (nts). Therefore the amino acid sequences corresponding to individual regions were aligned with the appropriate outgroups shown above using CLUSTAL X2 with TRANSALIGN (kindly supplied by Georg Weiller) to maintain the degapped alignment of the encoded amino acids and then the aligned subsequences were reassembled to form CP sequences 468 nt long. Putative recombination sites in all sequences were identified using RDP (Martin and Rybicki, 2000), GENECONV (Sawyer, 1999), BOOTSCAN (Salminen et al., 1995), MAXCHI (Smith, 1992), CHIMAERA (Posada and Crandall, 2001), and SISCAN programs (Gibbs et al., 2000) implemented in the RDP4 package (Martin et al., 2015) and also the original SISCAN version 2 (Gibbs et al., 2000) programs. First, we checked for incongruent relationships using the programs implemented in RDP4. These analyses were done using default settings for the different detection programs and a Bonferroni corrected *P*-value cut-off of 0.05 or 0.01, and then all isolates that had been identified as likely recombinants by the programs in RDP4, supported by three different methods with an associated *P*-value of $< 1.0 \times 10^{-6}$, were rechecked using the original SISCAN version 2. We checked 50 nucleotides slices of all sequences for evidence of recombination using these programs. These analyses also assessed which non-recombinant sequences had regions that were the closest to the regions of the recombinant sequences and hence indicated the likely lineages that provided those regions of the recombinant genomes. For convenience, we called them ‘parental isolates’ of recombinants. Having examined all sites with an associated *P*-value of $< 1 \times 10^{-6}$ (*i.e.*, the most likely recombination sites) (Tomitaka and Ohshima, 2006). The ORSV CP sequences were also aligned without an outgroup sequence and directly checked for evidence of recombination using the programs.

Phylogenetic analysis. The CP gene sequences of the ORSV isolates, together with that of BPMoV isolate, as outgroup, were aligned, as described above using CLUSTAL X2 with TRANSALIGN. The phylogenetic relationships of the aligned sequences were inferred by the maximum-likelihood method (ML) implemented in PhyML version 3 (Guindon and Gascuel, 2003). The best-fit model of nt substitutions for CP dataset was determined by jModeltest 0.1.1 (Posada, 2008); the general time-reversible substitution model with a gamma distribution of among-site rate variation (GTR + T4) provided the best fit for the data set. For ML tree, branch support was evaluated by the bootstrap method based on 1,000 pseudoreplicates. While the neighbor-joining (NJ) tree was also constructed

using MEGA version 7 (Kumar et al., 2016) with 1,000 bootstrap replicates. The calculated trees were displayed by TREEVIEW (Page, 1996).

Similarity, variation, and selection. The SDT version 1.2 software was used to estimate the pairwise nucleotide sequence identity scores (Muhire et al., 2014). The genetic distances (the average number of nt substitutions between two randomly selected sequences in a population) of the CP gene within and between geographic and phylogenetic groups were calculated also using Kimura's two-parameter correction method (Kimura, 1980) in MEGA version 7 (Kumar et al., 2016). Nonsynonymous (d_N) and synonymous (d_S) differences that correlated with phylogenetic relationships were accessed separately using the Pamilo–Bianchi–Li method (Pamilo and Bianchi, 1993) in MEGA version 7 (Kumar et al., 2016). Nucleotide and haplotype diversity of ORSV CP sequences were estimated by DnaSP version 6.0 (Rozas et al., 2017). Haplotype diversity was calculated based on the frequency and number of haplotypes in the population. Nt diversity estimates the average pairwise differences among sequences.

Genetic differentiation and gene flow. The three values of K_s^* , Z , and S_{nn} were used to examine the genetic differentiation between populations based on permutation statistical tests. To detect the level of gene flow between populations, the absolute value of the standardized variance in allele frequencies across populations (F_{st}) were estimated using DnaSP version 6.0 (Rozas et al., 2017). The absolute value of F_{st} ranges from 0 to 1 for undifferentiated to fully differentiated populations.

Estimation of the substitution rates, divergence times and spatial diffusion. BEAST v1.8.2 was used to estimate the evolutionary rate and timescale of the ORSV using the CP gene. Recombinant sequences were discarded. Bayes factors were used to select the best-fitting molecular-clock model. We compared strict and relaxed (uncorrelated exponential and uncorrelated lognormal) molecular clocks, as well as five demographic models (constant population size, expansion growth, exponential growth, logistic growth, and the Bayesian skyline plot). Posterior distributions of parameters, including the tree, were estimated by Markov chain Monte Carlo (MCMC) sampling. Samples were drawn every 10^4 MCMC steps over a total of 10^8 steps, with the first 10% of samples discarded as burn-in. Acceptable sampling from the posterior and convergence to the stationary distribution were checked using the diagnostic software Tracer version 1.6 (<http://tree.bio.ed.ac.uk/software/tracer/>). For

reliable rate estimates from time-structured sequence data, the spread of sampling times needs to be sufficient to allow an appreciable number of substitutions to occur. The temporal signal in our data sets was checked by comparing our rate estimates with those from 10 date-randomized replicates. A data set was considered to have an adequate spread in sampling times when its mean rate estimate was not contained in any of the 95% confidence intervals (CIs) of the rates estimated from replicates with randomized dates (Ramsden et al., 2009).

The spatial diffusion patterns of ORSV were inferred in BEAST using an asymmetric discrete-time Markov chain phylogeographic mode (Lemey et al., 2009). Seven geographical locations including mainland China, Germany, Australia, Singapore, South Korea, Indonesia, and Taiwan island, were selected and coded as discrete states. The migration pathways between pairs of locations can be identified by using Bayes factors > 10 and a mean indicator of > 0.5 (Suchard et al., 2001). The well-supported spatiotemporal movements of ORSV were identified using SPREAD v 1.0.7 (Bielejec et al., 2011) and Google Earth (<http://www.google.com/earth>).

Results

Nucleotide sequences. Twenty-one CP nt sequences of ORSV were determined here (Supplementary Table 1). In all the ORSV isolates sequenced in this study, the CP gene was 477 nt in length with 158 deduce aa. The CP sequences determined in this study are available in DDBJ/EMBL/GenBank databases with accession codes MG869629–MG869649. Three highly conserved RNA-binding motifs (I, II, and III) in tobamoviruses (Altschuh et al., 1987; Chng et al., 1996) were also observed from these ORSV CP gene sequences.

Recombination and phylogenetic analysis. Total 166 CP sequences obtained here and the public DNA sequence databases, were assessed for the evidence of recombination. After all gaps and nts homologous to them were removed from the aligned sequences, the likely recombination sites were assessed using RDP4 package and original SISCAN version 2 program. However, no recombination site was found in the CP gene region of all ORSV isolates. We calculated the ML and NJ trees based on the CP sequences of ORSV isolates, respectively. The ML tree is shown in Fig. 1. All ORSV isolates were clustered into one lineage, which is partially congruent with geography origin.

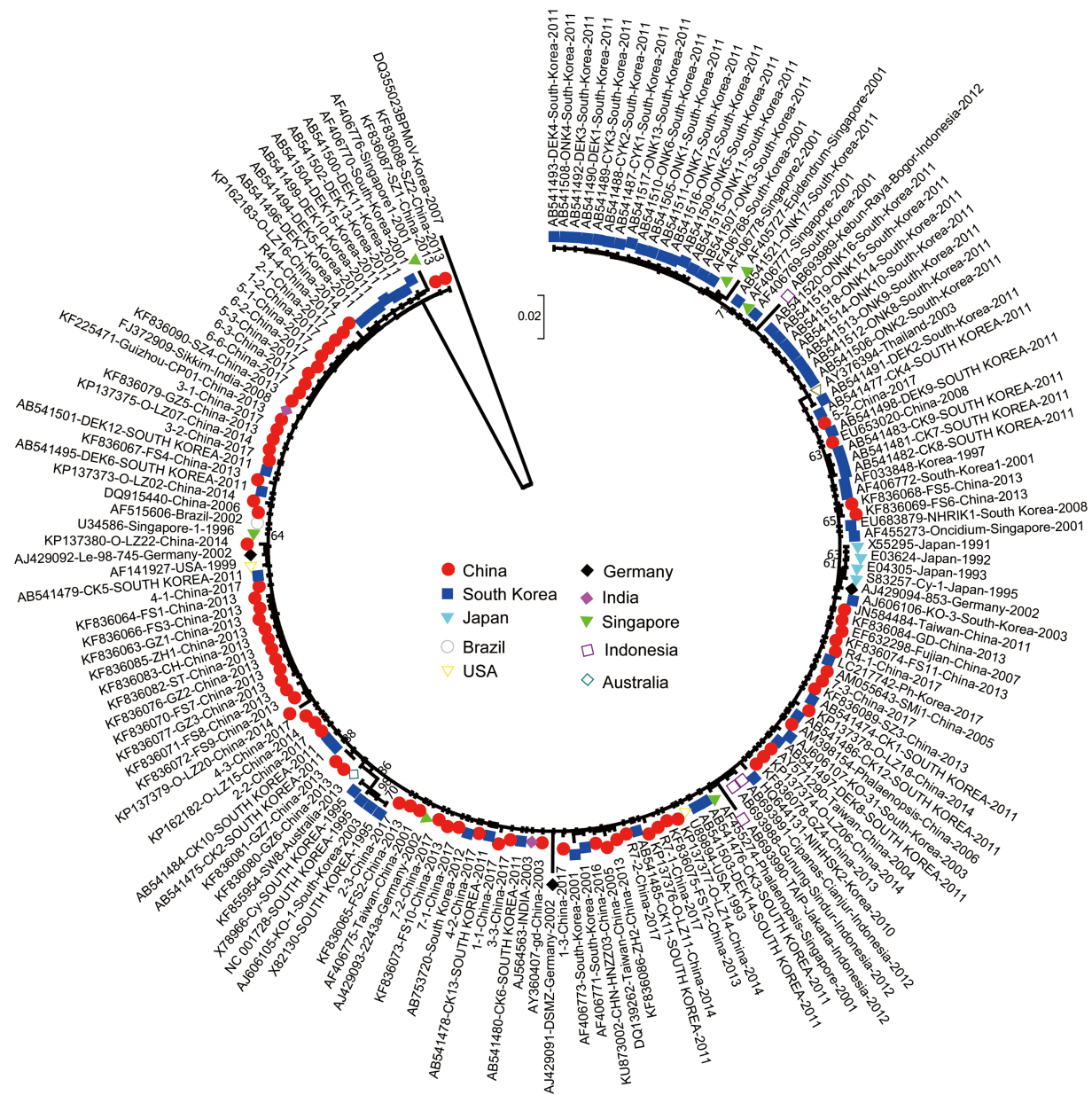


Fig. 1. Maximum-likelihood tree calculated from the coat protein gene sequences of *Odontoglossum ringspot virus* (ORSV). Numbers at each node indicate the percentage of supporting bootstrap samples in maximum-likelihood trees. The homologous sequences from an isolate of *Bell pepper mottle virus* (BPMoV) were used as outgroups.

Variation analysis. The similarities between the isolates in CP gene were 0.00823 (standard error, ± 0.00132) in nt sequences (Table 1). The nucleotide pairwise genetic identities of ORSV isolates in CP gene were calculated by SDT software. The lowest pairwise identity was approximately 94% (Supplementary Fig. 1). We also calculated the nt similarities within and between the Chinese, South Korean, German, Singaporean, Indonesian, and Japanese subpopulations (Table 2). The nt similarity for CP gene in

China showed the lowest values (0.00468) except in Japan where only four isolates were included in this study. The highest nt similarity was found in Indonesia (0.0252). The results suggest that there was small genetic differentiation in the CP gene between the geographic regions. Haplotype diversities of the geographic originated subpopulations were also compared (Table 1). In most cases, haplotype diversity values are high and nt diversity values are low except Japanese subpopulation.

Table 1. Nucleotide and haplotype diversities of the CP genes of *Odontoglossum ringspot virus* isolates

Group	d ^a	Hd ^b	d _N ^c	d _S ^c	d _N /d _S
All (n = 166)	0.00823 ± 0.000132	0.921 ± 0.017	0.00457 ± 0.000648	0.0146 ± 0.00351	0.31
China (n = 75)	0.00468 ± 0.000832	0.893 ± 0.033	0.00333 ± 0.000822	0.00747 ± 0.00179	0.45
South Korean (n = 65)	0.00826 ± 0.00207	0.894 ± 0.024	0.00329 ± 0.00104	0.0169 ± 0.00605	0.19
Germany (n = 4)	0.0185 ± 0.00446	1.000 ± 0.177	0.0112 ± 0.00394	0.0298 ± 0.0102	0.37
Singapore (n = 7)	0.0183 ± 0.00359	1.000 ± 0.076	0.0122 ± 0.00328	0.0272 ± 0.00898	0.45
Indonesia (n = 4)	0.0252 ± 0.00497	0.833 ± 0.222	0.0214 ± 0.00648	0.0302 ± 0.0104	0.71
Japan (n = 4)	0.00107 ± 0.00104	0.500 ± 0.265	0.00129 ± 0.00132	ND	ND

Values are presented as mean ± standard deviation.

CP, coat protein; ND, not determined.

^ad, nucleotide diversity estimated by the Kimura's two-parameter method.

^bHd, haplotype diversity.

^cd_N and d_S, nucleotide diversity at nonsynonymous and at synonymous position, respectively estimated by the Pamilo-Bianchi-Li method. The number (n) of sequences compared is indicated.

Table 2. With- and between-subpopulation nucleotide diversities in the coat protein gene region of *Odontoglossum ringspot virus*

Subpopulation 1	Subpopulation 2	Distance	Standard error
Chinese (n = 75)	German (n = 4)	0.0116509	0.0021736
Chinese (n = 75)	South Korean (n = 65)	0.0076695	0.0016644
German (n = 4)	South Korean (n = 65)	0.0145620	0.0026519
Chinese (n = 75)	Singaporean (n = 7)	0.0126104	0.0022931
German (n = 4)	Singaporean (n = 7)	0.0195863	0.0030364
South Korean (n = 65)	Singaporean (n = 7)	0.0131255	0.0025095
Chinese (n = 75)	Indonesian (n = 4)	0.0171066	0.0032910
German (n = 4)	Indonesian (n = 4)	0.0241356	0.0038720
South Korean (n = 65)	Indonesian (n = 4)	0.0187440	0.0036379
Singaporean (n = 7)	Indonesian (n = 4)	0.0235401	0.0039764
Chinese (n = 75)	Japanese (n = 4)	0.0072544	0.0029604
German (n = 4)	Japanese (n = 4)	0.0130477	0.0033453
South Korean (n = 65)	Japanese (n = 4)	0.0101344	0.0032954
Singaporean (n = 7)	Japanese (n = 4)	0.0151285	0.0035959
Indonesian (n = 4)	Japanese (n = 4)	0.0190868	0.0043660

Selection. The number of d_N was always smaller than d_S and varied considerably in Chinese, South Korean, German, Singaporean, Indonesian, and Japanese subpopulations. This suggests that there is selection against (*i.e.*, negative selection) most amino acid changes. We estimated the d_N/d_S ratios within the geographic originated subpopulations. The South Korean subpopulation had the highest selective constraints (0.1945), while the lowest ratios (0.7095) were found in Indonesia (Table 1). These results were consistent with the nt diversity analysis.

Genetic differentiation and gene flow. The K_s*, Z, and S_{nn} values of population differentiation supported that there has been partial genetic differentiation between Chinese, South Korean, German, Singaporean, Indonesian,

and Japanese subpopulations, consistently with the results of the phylogenetic analyses (Table 3).

The absolute values of *Fst* are used to analyze the level of gene flow between different geographical subpopulations (Table 3). In the present study, the absolute values of *Fst* were smaller than 0.33, suggesting that the gene flow of ORSV subpopulations among China, South Korea, Germany, Singapore, and Indonesia were frequent in CP gene. However, the absolute values of *Fst* for Japan with Chinese, South Korean, German, and Singaporean subpopulations were > 0.33, and this suggests that gene flow between the Japanese subpopulation and the other subpopulations are infrequent.

Evolutionary rates and timescales. We used a Bayesian

Table 3. Gene flow and genetic differentiation of *Odontoglossum ringspot virus* populations

Region 1	Region 2	Parameter ^a				
		<i>Ks</i> * (P-value)	<i>Z</i> (P-value)	<i>Snn</i> (P-value)	<i>Fst</i>	<i>Nm</i>
China (<i>n</i> = 75) ^b	Korea (<i>n</i> = 65)	2.94217 (0.0000***) ^c	4428.82278 (0.0000***)	0.74992 (0.0000***)	0.15573	1.36
	Germany (<i>n</i> = 4)	2.49641 (0.2820 ns)	1511.94313 (0.0360*)	0.94886 (0.0470*)	0.00470	52.90
	Singapore (<i>n</i> = 7)	2.70589 (0.0000***)	1598.08029 (0.0000***)	0.93951 (0.0000***)	0.09169	2.48
	Indonesia (<i>n</i> = 4)	2.64831 (0.0000***)	1490.84377 (0.0070**)	0.99367 (0.0000***)	0.12816	1.70
	Japan (<i>n</i> = 4)	2.09134 (0.0000***)	1439.05451 (0.0000***)	1.00000 (0.0000***)	0.60294	0.16
Korea (<i>n</i> = 65)	Germany (<i>n</i> = 4)	4.09692 (0.0720 ns)	1148.51979 (0.0480*)	0.94203 (0.0090**)	0.07909	2.91
	Singapore (<i>n</i> = 7)	4.26881 (0.4630 ns)	1257.30594 (0.0670 ns)	0.89331 (0.0010**)	-0.00911	-27.69
	Indonesia (<i>n</i> = 4)	4.27083 (0.0270*)	1131.93550 (0.0000***)	0.98551 (0.0000***)	0.10816	2.06
	Japan (<i>n</i> = 4)	3.63315 (0.0010**)	1115.73984 (0.0020**)	1.00000 (0.0000***)	0.53904	0.21
Germany (<i>n</i> = 4)	Singapore (<i>n</i> = 7)	8.42424 (0.2450 ns)	26.93197 (0.3650 ns)	0.58182 (0.2760 ns)	0.05843	4.03
	Indonesia (<i>n</i> = 4)	10.00000 (0.0310*)	12.75000 (0.0660 ns)	0.70833 (0.0310*)	0.09091	2.50
	Japan (<i>n</i> = 4)	4.50000 (0.0260*)	10.20833 (0.0260*)	0.87500 (0.0260*)	0.25000	0.75
Singapore (<i>n</i> = 7)	Indonesia (<i>n</i> = 4)	9.51515 (0.1790 ns)	25.12925 (0.1300 ns)	0.70606 (0.0700 ns)	0.07530	3.07
	Japan (<i>n</i> = 4)	5.51515 (0.0090**)	23.43537 (0.0380*)	1.00000 (0.0110*)	0.36239	0.44
Indonesia (<i>n</i> = 4)	Japan (<i>n</i> = 4)	6.00000 (0.0310*)	9.75000 (0.0310*)	1.00000 (0.0310*)	0.31429	0.55

^a*Ks** and *Z* are the sequence-based statistics. *Snn* is the nearest-neighbor statistic. *Fst* is the interpopulation component of genetic variation of the standardized variance in allele frequencies across populations. An absolute value of *Fst* < 0.33 suggests infrequent gene flow. *N* is the population size of each subpopulation. *m* is the migration fraction per generation.

^bThe number of sequences.

^c*P* < 0.05 was considered as the criterion for rejecting the null hypothesis that there is no genetic differentiation between two subpopulations.

*0.01 < *P* < 0.05; **0.001 < *P* < 0.01; ****P* < 0.001; ns, not significant.

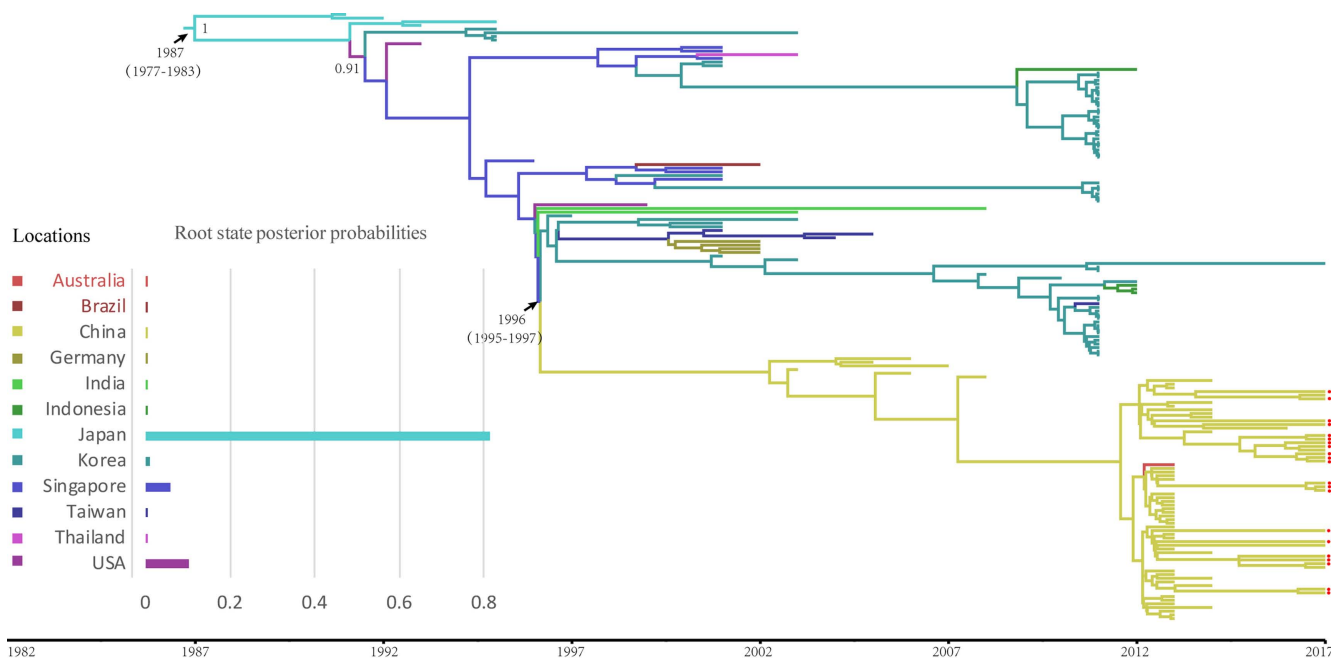


Fig. 2. Bayesian maximum-clade-credibility tree inferred from 166 sequences of the coat protein (CP) gene of *Odontoglossum ringspot virus*. The tree topology has been chosen to maximize the product of node posterior probabilities. Only posterior probability values above 0.90 are shown. The bar graph shows the root state posterior probabilities for each location. Isolates with red cycle are of Chinese origin and had their CP coding regions sequenced in this study. Year before present; 2017.

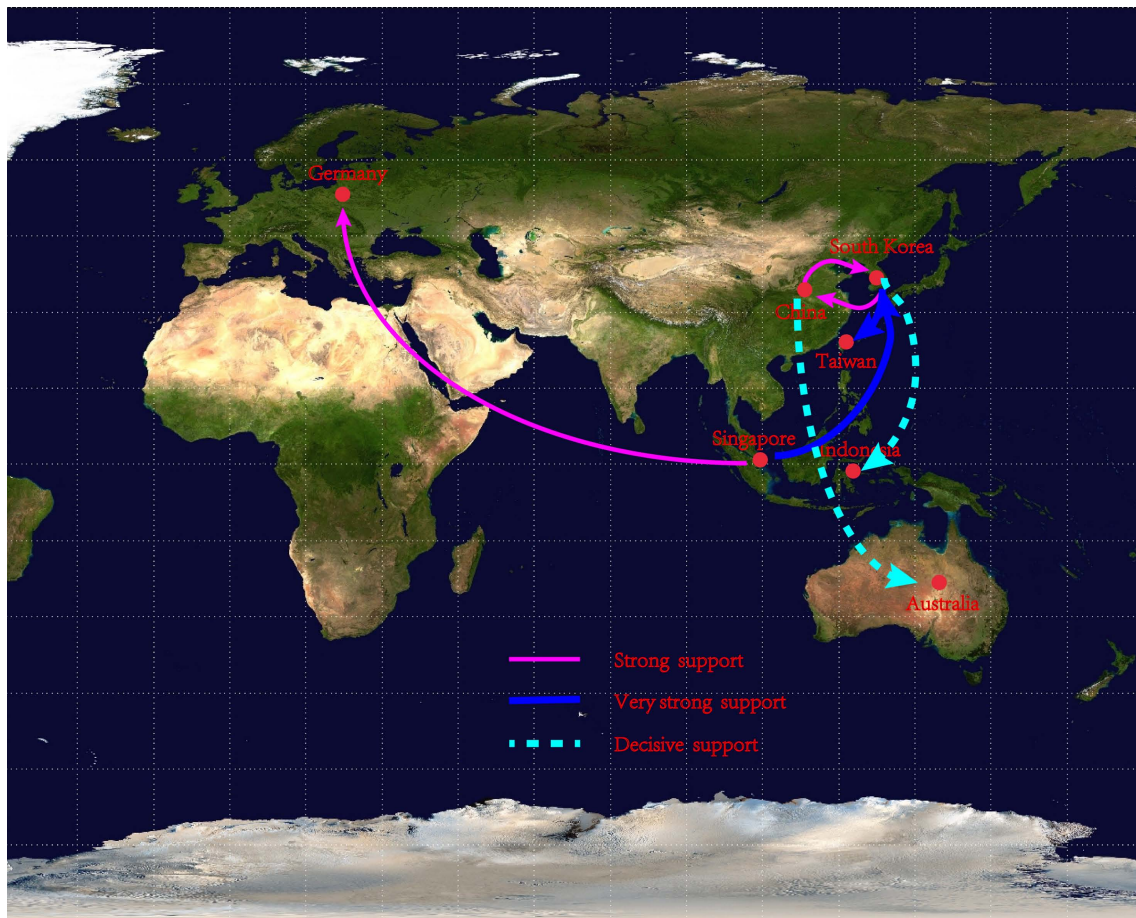


Fig. 3. Spatial diffusion pathways of *Odontoglossum ringspot virus* that were inferred using the coat protein gene sequences using non-recombinant sequences. Dissemination routes are only shown when supported by a Bayes factor (BF) > 10 and a mean indicator > 0.5. BF values can be interpreted as follows: $10 \leq \text{BF} < 30$, strong support; $30 \leq \text{BF} < 100$, very strong support; and $\text{BF} \geq 100$, decisive support.

phylogenetic method to estimate the evolutionary rates and timescales for ORSV CP gene. Based on a comparison of marginal likelihoods calculated using the harmonic-mean estimator in Tracer version 1.6 (<http://tree.bio.ed.ac.uk/software/tracer/>), the Bayesian skyline plot demographic model was the best support for ORSV. An uncorrelated lognormal relaxed-clock model provided a better fit than the strict-clock model, indicating the presence of rate variation among lineages. Our analyses of date-randomized replicates revealed that the sampling times of ORSV had sufficient temporal structure for calibration of the molecular clock (Supplementary Fig. 2). This was indicated by the smaller 95% credibility intervals of the rate estimates from the original data set compared with the date-randomized replicates (Ramsden et al., 2009). In addition, the mean posterior rate estimates from the original data were not contained within the 95% credibility intervals of the rate estimates from the date-randomized replicates (Ramsden et

al., 2009). The mean estimated substitution rates was 1.25×10^{-3} nt/site/y for CP. Estimates of the age of the root was 30 y (95% CI, 26-40) for CP (Fig. 2).

Geographical spread of ORSV. We assessed the likely migration routes of ORSV into mainland China, Germany, Australia, Singapore, South Korea, Indonesia, and Taiwan island according to the non-recombinant sequences of the CP gene using a Bayesian phylogeographical analysis (Lemey et al., 2009). Our results showed that ORSV migrated from east Asia countries (mainland China and South Korea) to Taiwan island, Germany, Australia, Singapore, and Indonesia (Fig. 3, Supplementary Table 2), and it also circulated within mainland China and South Korea, based on the CP data set. These results were confirmed by the gene flows and ML trees of the CP gene. The absolute values of *F_{st}* among mainland China, South Korea, Taiwan island, Singapore, and Indonesia

populations were all < 0.33 based on the CP gene (Table 3), which supported the frequent gene flows in east Asian, Germany, and Australia populations. In the CP ML tree, the Chinese and Korean isolates were closest to the German, Indonesian, Singaporean, and Australian isolates (Fig. 1), which also supported the probable migration routes.

Discussion

ORSV inducing viral diseases have been found in worldwide (Hu et al., 1993; Jensen and Gold, 1951; Pearson and Cole, 1986). The only effective control of ORSV is through the use of host plant genetic resistance, either conventional or transgenic (Ajjikuttira et al., 2002; Yoon et al., 2011). Hence, it is very important to understand the molecular variability, divergence, and migration of ORSV population in each region.

Recombination is an important source of genetic variation not only for potyviruses (Gibbs and Ohshima, 2010) but also for the viruses in the genus *Tobamovirus* (He et al., 2012; Lartey et al., 1996). However, the ORSV isolates in CP gene used in the study seemed to be non-recombinants, although previous report shown that homologous recombination had played a role in ORSV evolution (Lartey et al., 1996). The phylogenetic analyses together with the genetic differentiation analyses indicate that ORSV has a structure in China, South Korea, Germany, Singapore, Indonesia, and Japan. The varied d_N/d_S ratios obtained from Chinese, South Korean, German, Singaporean, Indonesian and Japanese subpopulations, also indicating that the ORSV CP gene is under different evolutionary constraints according to the geography. Analogously, there has been reported in several plant viruses where strong selection against amino acid change in the CP gene (Ge et al., 2014; He et al., 2014, 2016; Ogawa et al., 2008; Ohshima et al., 2016; Wei et al., 2009).

Recently, the evolutionary rates and the most recent common ancestors (TMRCAs) have been analyzed for *Tobacco mosaic virus* (TMV) (Gao et al., 2019), and flowers plant viruses (Ohshima et al., 2016). We estimated evolutionary rates and TMRCAs for ORSV. Ohshima et al. (2016) analyzed the evolutionary rates of five narcissus viruses were $1.33\text{--}7.15 \times 10^{-3}$ nt/site/y, while the four genes evolutionary rates of TMV were $3.92\text{--}4.38 \times 10^{-4}$ (Gao et al., 2019). Our analysis showed that ORSV had similar evolutionary rates with these narcissus viruses, but slightly faster than TMV. TMRCAs of ORSV was similar to *Narcissus degeneration virus* and TMV, but younger than *Narcissus late season yellows virus*, *Narcissus yellow stripe virus*, *Cyrtanthus elatus virus A*, and *Narcissus latent*

virus (Gao et al., 2019; Ohshima et al., 2016).

ORSV was first isolated from *Rossioglossum grande* in the early 1950s (Jensen and Gold, 1951). TMRCAs of ORSV (1977-1983) that was estimated here corresponds very closely to the timing of the virus emergence dates. The natural host range of ORSV is known to be limited to monocotyledonous orchid plants. Mark (2001) showed the family Orchidaceae is even older up to roughly 100 million years. Thus, our TMRCAs analysis indicated that the earlier ORSV populations were either not sampled or did not survive to the present day, which is similar to the recent estimate for *Potato virus S* (Duan et al., 2018).

In conclusion, the CP gene of ORSV is evolving at a rate of 1.25×10^{-3} nt/site/y, which is similar to that of RNA and ssDNA viruses. TMRCAs of CP gene was 30 y (95% CIs, 26-40). ORSV migrated from east Asia countries (mainland China and South Korea) to Taiwan island, Germany, Australia, Singapore, and Indonesia, and it also circulated within mainland China and South Korea. Our analysis of ORSV isolates from worldwide countries provides useful information about the evolution and genetic conservation of the viruses in the genus of *Tobamovirus*.

Acknowledgments

This work was supported by grants from the Key project at central government level: The ability establishment of sustainable use for valuable Chinese medicine resources (2060302), National Natural Science Foundation of China (No. 31601604) and China Agriculture Research System (CARS-24).

Electronic Supplementary Material

Supplementary materials are available at The Plant Pathology Journal website (<http://www.ppjonline.org/>).

References

- Ajjikuttira, P. A., Lim-Ho, C. L., Woon, M. H., Ryu, K. H., Chang, C. A., Loh, C. S. and Wong, S. M. 2002. Genetic variability in the coat protein genes of two orchid viruses: *Cymbidium mosaic virus* and *Odontoglossum ringspot virus*. *Arch Virol.* 147:1943-1954.
- Altschuh, D., Lesk, A. M., Bloomer, A. C. and Klug, A. 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to *Tobacco mosaic virus*. *J Mol Biol.* 193:693-707.
- Bielejec, F., Rambaut, A., Suchard, M. A. and Lemey, P. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* 27:2910-2912.

- Chng, C.-G., Wong, S.-M., Mahtani, P. H., Loh, C.-S., Goh, C.-J., Kao, M. C.-C., Chung, M. C.-M. and Watanabe, Y. 1996. The complete sequence of a Singapore isolate of *Odontoglossum ringspot virus* and comparison with other tobamoviruses. *Gene* 171:155-161.
- Domingo, E. and Holland, J. J. 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51:151-178.
- Duan, G., Zhan, F., Du, Z., Ho, S. Y. W. and Gao, F. 2018. Europe was a hub for the global spread of *Potato virus S* in the 19th century. *Virology* 525:200-204.
- García-Arenal, F., Fraile, A. and Malpica, J. M. 2001. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* 39:157-186.
- Gao, F., Liu, X., Du, Z., Hou, H., Wang, X., Wang, F. and Yang, J. 2019. Bayesian phylodynamic analysis reveals the dispersal patterns of *Tobacco mosaic virus* in China. *Virology* 528:110-117.
- Ge, B., He, Z., Zhang, Z., Wang, H. and Li, S. 2014. Genetic variation in *Potato virus M* isolates infecting pepino (*Solanum muricatum*) in China. *Arch. Virol.* 159:3197-3210.
- Gibbs, A. and Ohshima, K. 2010. Potyviruses and the digital revolution. *Annu. Rev. Phytopathol.* 48:205-223.
- Gibbs, M. J., Armstrong, J. S. and Gibbs, A. J. 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16:573-582.
- Guindon, S. and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696-704.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95-98.
- He, M., He, C.-Q. and Ding, N.-Z. 2012. Natural recombination between tobacco and tomato mosaic viruses. *Virus Res.* 163:374-379.
- He, Z., Li, W., Yasaka, R., Huang, Y., Zhang, Z., Ohshima, K. and Li, S. 2014. Molecular variability of *Sugarcane streak mosaic virus* in China based on an analysis of the P1 and CP protein coding regions. *Arch. Virol.* 159:1149-1154.
- He, Z., Yasaka, R., Li, W., Li, S. and Ohshima, K. 2016. Genetic structure of populations of *Sugarcane streak mosaic virus* in China: comparison with the populations in India. *Virus Res.* 211:103-116.
- Hu, J. S., Ferreira, S., Wang, M. and Xu, M. Q. 1993. Detection of *Cymbidium mosaic virus*, *Odontoglossum ringspot virus*, *Tomato spotted wilt virus*, and potyviruses infecting orchids in Hawaii. *Plant Dis.* 77:464-468.
- Jensen, D. D. and Gold, A. H. 1951. A virus ringspot of *Odontoglossum* orchid: symptoms, transmission, and electron microscopy. *Phytopathology* 41:648-653.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-120.
- Kumar, S., Stecher, G. and Tamura, K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33:1870-1874.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-2948.
- Lartey, R. T., Voss, T. C. and Melcher, U. 1996. *Tobamovirus* evolution: gene overlaps, recombination, and taxonomic implications. *Mol. Biol. Evol.* 13:1327-1338.
- Lemey, P., Rambaut, A., Drummond, A. J. and Suchard, M. A. 2009. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5:e1000520.
- Mark, W. C. 2001. The origin and biogeography of *Orchidaceae*. In: *Genera Orchidacearum. Vol. 2. Orchidoideae (Part 1)*, eds. by A. M. Pridgeon, P. J. Cribb, M. W. Chase and F. Rasmussen, pp. 1-5. Oxford University Press, Oxford, UK.
- Martin, D. P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1:vev003.
- Martin, D. and Rybicki, E. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16:562-563.
- Muhire, B. M., Varsani, A. and Martin, D. P. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* 9:e108277.
- Ogawa, T., Tomitaka, Y., Nakagawa, A. and Ohshima, K. 2008. Genetic structure of a population of *Potato virus Y* inducing potato tuber necrotic ringspot disease in Japan: comparison with North American and European populations. *Virus Res.* 131:199-212.
- Ohshima, K., Nomiya, R., Mitoma, S., Honda, Y., Yasaka, R. and Tomimura, K. 2016. Evolutionary rates and genetic diversities of mixed potyviruses in *Narcissus*. *Infect. Genet. Evol.* 45:213-223.
- Page, R. D. M. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Bioinformatics* 12:357-358.
- Pamilo, P. and Bianchi, N. O. 1993. Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. *Mol. Biol. Evol.* 10:271-281.
- Pearson, M. N. and Cole, J. S. 1986. The effects of *Cymbidium mosaic virus* and *Odontoglossum ringspot virus* on the growth of cymbidium orchids. *J. Phytopathol.* 117:193-197.
- Posada, D. 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25:1253-1256.
- Posada, D. and Crandall, K. A. 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc. Natl. Acad. Sci. U. S. A.* 98:13757-13762.
- Ramsden, C., Holmes, E. C. and Charleston, M. A. 2009. Hanta-virus evolution in relation to its rodent and insectivore hosts: no evidence for codivergence. *Mol. Biol. Evol.* 26:143-153.
- Rao, X., Li, Y., Sun, J., Li, X., Li, M. and Xiang, M. 2015. Genetic diversities of *Cymbidium mosaic virus* and *Odontoglossum ringspot virus* isolates based on the coat protein genes from orchids in Guangdong province, China. *J. Phytopathol.*

- 163:324-329.
- Rhie, M. J., Min, B. E., Hong, J. S., Song, Y. S. and Ryu, K. H. 2007. Complete genome sequence supports *Bell pepper motile virus* as a species of the genus *Tobamovirus*. *Arch. Virol.* 152:1401-1407.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. and Sánchez-Gracia, A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34:3299-3302.
- Ryu, K. H. and Park, W. M. 1995. The complete nucleotide sequence and genome organization of *Odontoglossum ringspot tobamovirus* RNA. *Arch. Virol.* 140:1577-1587.
- Salminen, M. O., Carr, J. K., Burke, D. S. and McCutchan, F. E. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res. Hum. Retroviruses* 11:1423-1425.
- Sawyer, S. A. 1999. GENECONV: a computer package for the statistical detection of gene conversion. Distributed by the author, Department of Mathematics, Washington University in Louis. URL <http://www.math.wustl.edu/sawyer> [22 April 2019].
- Smith, J. M. 1992. Analyzing the mosaic structure of genes. *J. Mol. Evol.* 34:126-129.
- Suchard, M. A., Weiss, R. E. and Sinsheimer, J. S. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* 18:1001-1013.
- Tomitaka, Y. and Ohshima, K. 2006. A phylogeographical study of the *Turnip mosaic virus* population in East Asia reveals an “emergent” lineage in Japan. *Mol. Ecol.* 15:4437-4457.
- Wei, T. Y., Yang, J. G., Liao, F. L., Gao, F. L., Lu, L. M., Zhang, X. T., Li, F., Wu, Z. J., Lin, Q. Y., Xie, L. H. and Lin, H. X. 2009. Genetic diversity and population structure of *Rice stripe virus* in China. *J. Gen. Virol.* 90:1025-1034.
- Wong, S. M., Chng, C. G., Lee, Y. H., Tan, K. and Zettler, F. W. 1994. Incidence of *Cymbidium mosaic* and *odontoglossum ringspot* viruses and their significance in orchid cultivation in Singapore. *Crop Prot.* 13:235-239.
- Yoon, J.-Y., Chung, B.-N. and Choi, S.-K. 2011. High sequence conservation among *Odontoglossum ringspot virus* isolates from orchids. *Virus Genes* 42:261-267.
- Zettler, F. W., Ko, N. J., Wisler, G. C., Elliott, M. S. and Wong, S. M. 1990. Viruses of orchids and their control. *Plant Dis.* 74:621-626.
- Zheng, Y. X., Shen, B. N., Chen, C. C. and Jan, F. J. 2010. *Odontoglossum ringspot virus* causing flower crinkle in *Phalaenopsis* hybrids. *Eur. J. Plant Pathol.* 128:1-5.