Eur J Plant Pathol (2011) 129:361–370
DOI 10.1007/s10658-010-9716-8

Occurrence of Honeysuckle Yellow Vein Virus (HYVV) containing a monopartite DNA-A genome in Korea

Yuan Wang · Jing Ji · Tae-Kyun Oh · Sung Oh · Sue Hoon Kim · Hyun Ju Lee · Myoung Yong Shim · Chang Won Choi · Seong Hwan Kim · Il-Seop Kim · Young Shik Kim

Accepted: 15 November 2010 / Published online: 23 December 2010

Abstract Two newly emerged begomoviruses were isolated from naturally infected tomato (*Solanum lycopersicum*) plants grown in greenhouses at Jeju Island and Dangjin in Korea and their genomes were characterized. These viruses-infected plants had very small leaves that curled upward, yellow margins and a leathery appearance, and a bushy and stunted appearance with short internodes. Nucleotide (nt) sequence analysis of their genomes showed that they have a DNA-A component of a monopartite begomovirus. Their genomes comprised 2763 and 2764 nucleotides with six open reading frames. The results of nt sequence similarity analysis of DNA-A genome between the two Korean isolates and isolates of *Tobacco leaf curl Japan virus* (TbLCJV), *Honeysuckle yellow vein virus* (HYVV), *Honeysuckle yellow vein mosaic virus* (HYVMV), and Eupatorium *yellow vein virus* in Japan (EpYVV) showed that they are likely similar to HYVV-[Masuda] (89.4–92.8% nt identity). Consequently, we tentatively propose the two isolates' names as HYVV-Jeju and -DJ according to the ICTV geminivirus rules. Phylogenetic relationship analysis of 33 DNA-A genome sequences using PAUP* 4.0b10 and MrBayes revealed that HYVV-Jeju and -DJ belong to the Far East Asian begomovirus species complex. Within the Far East Asian begomovirus species complex, HYVV-Jeju and -DJ are distantly related to EpYVV, HYVMV, and TbLCJV groups. Based on the presence of a recombination fragment spanning the C3 ORF, a recombinant origin was suggested for both HYVV-Jeju and –DJ, with parents close to Japanese isolates HYVMV-[SP1:00] and *Eupatorium yellow vein virus* (EpYVV)-[Suya]. In addition, the presence of a further recombination fragment spanning the IR suggested the parents of HYVV-DJ were close to HYVV-Jeju and EpYVV-[Suya].

Keywords Begomovirus · Tomato · DNA-A · HYVV-Jeju · HYVV-DJ · Recombination

Yuan Wang, Jing Ji, Tae-Kyun Oh and Sung Oh contributed equally to this study.

Y. Wang · J. Ji · T.-K. Oh · S. Oh · S. H. Kim · H. J. Lee · M. Y. Shim · C. W. Choi (✉)
Department of Biology and Medicinal Science,
Pai Chai University,
Daejeon 302-735, Korea
e-mail: choicw@pcu.ac.kr

S. H. Kim
Department of Microbiology and Institute of Basic Sciences, Dankook University,
Cheonan, Chungnam 330-714, Korea

I.-S. Kim
Department of Horticulture Science,
Kangwon National University,
Chuncheon, Gangwon-do 200-701, Korea

Y. S. Kim
Deptartment of Plant Science and Technology,
Sangmyung University,
Cheonan, Chungnam 330-720, Korea

Some species in the *Begomovirus* genus of the *Geminiviridae* family transmitted by the whitefly (*Bemisia tabaci*) have a circular single-stranded DNA genome of approximately 2700 nucleotides (nt). The genome has six partially overlapping open reading frames (ORFs) that are organized bi-directionally; two of these ORFs (V1 and V2) are in the virion sense orientation, and four (C1–C4) in the complementary orientation. Between the two transcription units lies an intergenic region (IR) which has key factors for the replication and transcription of the viral genome and is organized in a typical iterative structure. The majority of begomoviruses appear to have bipartite genomes of A and B DNA components (Lazarowitz 1992). DNA A is essential for replication, regulation of gene expression and encapsidation, while DNA B is responsible for viral movement between plant tissues, host range and symptom modulation (Hanley-Bowdoin et al. 2000). On the other hand, DNA B is lacking in some begomoviruses (Ueda et al. 2004; Pietersen et al. 2008).

To date, four phylogenetically-related begomoviruses have been reported in Japan. *Tobacco leaf curl Japan virus* (TbLCJV) causes yellow dwarf symptom in tomato and curling disease in tobacco (Shimizu and Ikegami 1999). *Eupatorium yellow vein virus* (EpYVV) causes yellow vein symptom in *Eupatorium makinoi* as a host plant in nature (Yahara et al. 1995). *Honeysuckle yellow vein virus* (HYVV) causes yellow net symptoms accompanied by small elliptical enations on the underside of the leaves of *Lonicera japonica* (honeysuckle), while *Honeysuckle yellow vein mosaic virus* (HYVMV) causes yellowing of veins followed by mosaic symptoms in honeysuckle, known as the natural overwintering host of TbLCJV (Ogawa et al. 2008). Though HYVMV is a begomovirus which is distinct from TbLCJV (Fauquet and Stanley 2005), pairwise comparison analyses of the nt sequences of DNA-A revealed that these viruses shared a significant threshold level of 84% identity. It suggested that TbLCJV, EpYVV and HYVMV might have a common ancestor and diverged through adaptation to *Eupatorium* and honeysuckle plants (Ueda et al. 2008). In this study, we have characterized two newly emerged tomato-infecting begomoviruses in Korea by determining the nt sequences of their complete genomes and the deduced amino acid (aa) sequences. Here we report the two Korean isolates are a new type of HYVV that belong to a Far East Asian begomovirus group and distantly related to known HYVV isolates.

Three virus-infected tomato plants showing begomovirus-like symptoms were collected from commercial greenhouses of Jeju-si, Jeju Island (July 2007) and Dangjin, Chungnam Province (Aug 2009) in Korea. Jeju Island (33n31, 126e32) is located at the southernmost part of Korean peninsula and its climate is subtropical with an average annual temperature of 15.5°C. Dangjin (36n54, 126e39) is located at the central-west coast of Korean peninsula and its climate is temperate (continental) with an average annual temperature of 11.4°C. Infected plants had very small leaves that curled upward, with yellow margins and a leathery appearance, and developed a bushy, stunted appearance with short internodes (Fig. 1). Emergence of a tomato-infecting virus in the local area has been related to the spread of its vector, the whitefly *B. tabaci*. This is the first characterization of newly emerged tomato disease caused by begomoviruses in tomato from Korea. Although any nationwide epidemiological study has not been conducted yet, these viruses seem to be spreading to other areas in Korea.

Total genomic DNA was isolated from 100 mg of leaf tissues of the virus-infected plant using the DNeasy Plant Mini Kit (QIAGEN, Germany) according to the instructions of the manufacturer. The DNA extract was used for amplification of viral genome

Fig. 1 Tomato foliar symptoms (*arrows*) showing curling and yellowing on the edges, caused by HYVV-Jeju isolated from Jeju Island in Korea

fragments by PCR using the following forward and reverse primers according to the conserved nt sequences in the same region of numerous begomoviruses in the NCBI database listed in Table 1 (nt 2591-340 covering C1-V1: 5′-GAATTCAAAACGTTTCGGTGG-3′ and 5′-GAATTCAGACGGGAGTGAAAA-3′, nt 298–1071 covering V1: 5′-GAATTCATGTCGAA GCGWCCA-3′ and 5′-GAATTCTTAATTTKRTAYT GAATCATAGAA-3′, nt 1068–1260 covering V1-C1: 5′-GAATTCGATTCACGCACAGGG-3′ and 5′- GTC GACTTAAATACCCTCAAGAAATG-3′, nt 1520–2611 covering C1: 5′-ATGGCTCCACCGAAACGTTT TAGAATA-3′ and 5′- TCAATCCTCCTCCTGCG TGGTTGT-3′, nt 1–2763(4) covering complete genome: 5′-ATTACCGGATGGCCGCCCGAA-3′ and 5′-ATTA TACGGATGGCCGCTTTACCAATT-3′). The reaction mixture for PCR included 5 μl of 10×PCR buffer (100 mM Tris–HCl, pH 8.3, 15 mM $MgCl_2$ and 500 mM KCl), 4 μl of 2.5 mM deoxynucleotide triphosphates each, 1 μl of each forward and reverse primer (stock concentration, 50 ρM), 0.5 μl (2.5U) of Ex-*Taq* DNA polymerase (Takara, Japan), 2 μl of template genomic DNA and sterile distilled water which was added to give a 50 μl final volume. PCR was carried out in a thermal cycler (Bio-Rad) with initial denaturation at 95°C for 2 min, 30 cycles of denaturation at 94°C for 1 min, annealing at 50°C (C1-V1), 52°C (V1), 54°C (V1-C1), 58°C (C1) and 58°C (complete genome) for 2 min, and extension at 72°C for 3 min. The PCR-amplified viral gene fragments were subsequently cloned into pGEM-T easy vector (Promega, USA) using the manufacturer's protocol. Purification of PCR products was done using a PCR purification kit (iNtRON Biotechnology, Korea) according to the protocol of the manufacturer. The ligation mixture was transformed into DH5α *Escherichia coli* competent cells by heat shock method (Froger and Hall 2007). The transformed colonies were selected on Luria Bertani (LB)-ampicillin agar media containing IPTG and X-gal. The selected colonies were grown in LB liquid media and used for plasmid isolation using a commercial plasmid purification kit (iNtRON, Korea). The presence of the insert DNA in the plasmid was verified using restriction enzyme digestion with *Not*I (Takara, Japan). The pGEM-T easy vector containing the target insert DNA was sent to a commercial sequencing lab (COSMOgenetech, Korea) to determine its nt sequences and deduced aa sequences. At least three independent clones were sequenced for each amplicon from each tomato sample collected in the same greenhouse. For sequence comparisons, the determined nt and aa sequences were aligned with those of reference begomovirus isolates and other geminiviruses obtained from the NCBI database. The identities of nt and aa sequences were determined using the DNAMAN software version 7.0 (Lynnon Biosoft, Quebec, Canada).

Identical nt sequences were obtained from all clones containing the full-length or partial-length genome of HYVV-Jeju and -DJ isolates, respectively, from individual leaf samples of three tomato plants of each location. This result indicated that three tomato plants collected from each location were infected with the same virus. The nt sequence data have been deposited to the GenBank database as accession numbers FJ434943 (HYVV-Jeju) and HQ189431 (HYVV-DJ). The genomes of HYVV-Jeju and -DJ comprised 2763 nt and 2764, respectively, with six ORFs (2 ORFs on the virion strand and 4 ORFs on the complementary strand, the typical genome organization of begomoviruses originating from the Old World). In recent years, a satellite DNA, referred to as DNAβ, was reported to be associated with EpYVV (Saunders et al. 2000), TbLCJV (Ogawa et al. 2007), and HYVMV (Were et al. 2005). Nucleotide sequences of DNAβ satellite molecules associated with three viruses were determined in Japan (Ogawa et al. 2008; Ueda et al. 2008). Therefore, we tried to detect the presence of a DNAβ satellite from HYVV-Jeju and -DJ infected tomato leaves using known primers under various conditions of PCR reaction, as described in previous studies. However, we failed to detect the DNAβ satellite molecule. On the basis of negative results, we suspect that two HYVV isolates found in Korea may be satellite-free isolates. However, further efforts to identify the DNAβ satellite molecule are currently underway.

The full genome sequences of HYVV-Jeju and -DJ were compared to numerous begomovirus sequences in Table 1. HYVV-Jeju exhibited the highest nt sequence identity to genome sequences of the following begomovirus isolates: HYVV-[Masuda] (92.8%) and TbLCJV-[Koc] (92.6%, previously known as *Tobacco leaf curl Kochi virus*). HYVV-DJ exhibited the highest nt sequence identity to genome sequences of the HYVV-Jeju (91.9%). Begomoviruses sharing less than 89% sequence identity in DNA-A are considered to be distinct species, while the viruses sharing more than

Table 1 Percentage nucleotide (nt) and deduced amino acid (aa) sequence identities between HYVV-Jeju and other begomovirus isolates

Virus isolates[a]	Total nt	IR nt	V1 nt/aa	V2 nt/aa	C1 nt/aa	C2 nt/aa	C3 nt/aa	C4 nt/aa
HYVV-[Masuda]	92.8	90.4	92.2/96.9	88.0/89.7	95.1/84.9	91.9/89.6	91.9/86.6	96.6/92.8
TbLCJV-[Koc]	92.6	89.7	93.9/98.4	93.3/94.7	93.4/84.0	92.9/90.4	92.1/90.3	93.5/83.5
HYVMV-[HY12]	90.5	85.9	90.1/93.4	86.6/85.3	92.9/82.6	91.9/88.2	89.6/85.8	94.9/86.6
TbLJV-[Iba:06]	90.4	84.1	93.4/97.3	87.5/89.7	91.4/92.0	89.0/85.2	89.4/86.6	93.2/86.6
TbLCJV-[JP3]	89.8	87.3	92.6/98.1	86.8/86.2	91.7/82.0	88.5/85.2	89.6/84.3	96.3/90.0
HYVMV-[FK1]	90.2	89.3	85.4/86.0	86.9/88.8	94.6/92.6	88.2/84.4	88.6/85.1	96.9/93.8
TbLCJV-2007[b]	89.7	85.6	92.9/96.9	87.5/87.9	91.4/82.0	89.0/85.2	88.6/85.1	93.5/87.7
TbLCJV-1999 [b]	87.1	68.8	91.3/96.1	88.9/87.9	89.9/90.9	88.7/85.2	88.9/83.6	93.2/83.5
HYVMV-[OT2]	90.3	88.9	85.5/86.8	86.6/88.8	94.8/82.4	88.2/83.7	88.9/85.1	96.9/93.8
TbLCJV-[JP2]	89.4	86.6	91.2/96.1	86.9/87.1	91.5/81.8	88.5/85.2	88.9/82.8	95.6/87.6
HYVMV-[OT1]	89.8	89.7	85.4/86.0	86.3/88.8	94.1/92.8	88.2/82.2	87.9/83.6	97.3/93.8
HYVV-[HYVVNZ1]	87.5	59.2	90.1/94.2	87.2/87.9	90.1/88.7	92.2/88.2	90.1/86.6	92.2/83.5
HYVV-[Fukui]	87.8	59.9	91.3/97.7	86.9/85.3	83.3/72.5	91.7/88.2	90.1/85.1	91.8/82.5
HYVMV-[NarI:06]	89.5	87.2	85.9/88.0	86.3/86.2	92.7/82.4	89.7/85.2	89.9/85.8	93.3/85.9
HYVMV–[TobKG5]	88.7	85.1	85.5/85.2	84.3/85.3	92.1/92.6	88.5/84.4	89.4/85.1	94.2/86.6
HYVMV-[Yam]	87.2	67.4	86.2/86.8	86.0/87.9	91.8/93.1	91.7/88.2	90.6/86.6	92.2/82.5
HYVMV-[Miyazaki]	86.8	67.8	86.6/88.7	86.9/87.1	91.8/91.2	89.0/83.7	88.2/82.1	66.0/62.9
HYVMV-[Nar2:06]	86.4	68.5	85.5/87.9	85.2/84.5	90.7/91.5	90.7/85.9	89.1/85.1	93.2/83.5
HYVMV-[SP1:00]	85.8	70.5	85.3/87.6	84.3/80.2	88.9/76.9	90.2/85.2	89.0/77.8	91.2/80.4
EpYVV-[Suya]	81.6	52.4	85.5/87.9	83.8/84.5	85.8/78.8	82.0/69.1	84.2/79.1	91.2/79.4
EpYVV-[Kagawa]	80.6	49.5	85.5/88.7	84.6/84.5	85.6/88.2	82.8/71.1	85.4/79.1	91.2/80.4
EpYVV-[Yam]	80.5	48.0	85.8/88.7	84.6/83.6	85.3/86.5	82.4/71.9	85.2/79.9	62.4/55.7
EpYVV-[MNS2]	80.5	46.2	84.9/88.3	82.9/. 84.5	85.4/86.8	82.8/71.1	85.4/79.1	62.4/55.7
EpYVV-[SOJ3]	79.0	63.9	84.8/87.9	83.2/86.2	79.3/81.8	84.1/72.6	85.2/79.9	67.6/44.1
HYVV-[UK1]	86.5	57.4	91.2/96.1	86.9/87.1	83.0/81.4	91.2/86.7	89.1/85.1	49.3/44.3
HYVV-[UK2]	87.1	58.4	90.6/94.6	86.6/86.2	89.8/88.2	91.7/88.2	89.6/85.8	91.8/82.5
SgYVGdV	76.3	49.8	74.8/79.7	72.7/62.5	82.6/85.4	76.4/61.8	80.0/75.4	84.4/68.0
PaLCuCNV	72.7	43.1	75.6/79.8	76.9/75.0	83.1/85.1	72.6/60.7	73.8/68.7	86.4/71.1
TbLCZWV	69.0	43.7	69.8/72.5	69.2/63.8	75.0/74.9	67.7/52.6	72.8/67.2	68.4/47.4
TbLCYV-[Y136]	73.1	46.0	74.6/73.9	71.4/68.9	80.4/68.6	71.8/59.3	73.1/66.4	74.2/52.6
TbLCThV	72.8	44.8	75.2/75.5	72.2/71.4	81.0/84.0	72.8/60.0	73.6/70.2	84.0/70.1
HYVV-DJ	91.9	70.2	95.2/98.1	98.3/97.4	92.6/94.5	95.8/94.8	97.3/97.0	93.9/85.6

[a] GenBank accession numbers: HYVV-Jeju (FJ434943), HYVV-[Masuda] (AB236325), TbLCJV-[Koc] (AB055009), HYVMV-[HY12] (AB178946), TbLJV-[Iba:06] (AB287439), TbLCJV-[JP3] (AB079689), HYVMV-[FK1] (AB178945), TbLCJV-2007 (EF620536), TbLCJV-1999 (AB028604), HYVMV-[OT2] (AB178948), TbLCJV-[JP2] (AB055008), HYVMV-[OT1] (AB178947), HYVV-[HYVVNZ1] (FJ817425), HYVV-[Fukui] (AB236321), HYVMV-[NarI06] (AB287440), HYVMV–[TobKG5] (AB178949), HYVMV-[Yam] (AB079765), HYVMV-[Miyazaki] (AB236323), HYVMV-[Nar2:06] (AB287441), HYVMV-[SP100] (AB182261), EpYVV-[Suya] (AB300463), EpYVV-[Kagawa] (AB433979), EpYVV-[Yam] (AB079766), EpYVV-[MNS2] (AJ438936), EpYVV-[SOJ3] (AJ438937), HYVV-[UK1] (AJ542540), HYVV-[UK2] (AJ543429), SgYVGdV (AM238692), PaLCuCNV (AJ876548), TbLCZWV (AF350330), TbLCYV-[Y136] (AJ512761.1), TbLCThV (DQ871221) and HYVV-DJ (HQ189431)

[b] *Tobacco leaf curl Japan virus* (EF620536) and *Tobacco leaf curl Japan virus* (AB028604) are described as TbLCJV-2007 and -1999 in this study

89% are considered to belong to the one species (Fauquet et al. 2008). On the basis of the guideline proposed in geminivirus nomenclature (Fauquet et al. 2008), HYVV-Jeju and -DJ are considered the isolates of HYVV. When nt sequences of the IR of HYVV-Jeju were also compared with those of other begomoviruses in Table 1, the highest sequence identity was found with that of HYVV-[Masuda] (90.4%), while the lowest sequence identity was found with that of *Papaya leaf curl China virus* (PaLCuCNV, 43.1%). The IR of HYVV-DJ shared the highest sequence identity with that of TbLCJV-1999 (84.8%), while it shared the lowest sequence identity with that of TbLCThV (50.5%).

In HYVV-Jeju (-DJ), a 33-base potential stem-loop forming region (nt coordinates 2744(5)–2764 (5) and complementary 1–13: 5′-GCGGCCATCCG TATAATATTACCGGATGGCCGC-3′) was found in a 289 nt IR between the C1 and V2 ORFs. It included the nonanucleotide motif TAATATTAC (nt coordinates 2757(8)-02) presenting at the apex of a potential stem-loop structure (Lazarowitz 1992), a conserved begomovirus sequence involved in binding to the replication-associated protein (Rep). Other inverted repeats, known as putative Rep-binding sites, are found at nt coordinates 2630–2634 of HYVV-Jeju and 2627–2631 of HYVV-DJ (5′-GTACC-3′) and nt coordinates 2679–2683 (2680–2684) (5′-GGTAC-3′). In the IR, there are eukaryotic promoter TATA boxes (TATATA) oriented in both senses (nt coordinates 105–110 and 2686–2691 (105–110 and 2670–2675).

In individual encoded proteins of both HYVV-Jeju and -DJ were V1: 257 aa, 298–1071 nt, V2: 113 aa, 138–479 nt, C1: 363 aa, 1520–2611 nt, C2: 135 aa, 1213–1620 nt, C3: 134 aa, 1068–1472 nt and C4: 97 aa, 2161–2454 nt. HYVV-Jeju shared the highest aa sequence identities with TbLCJV-[Koc] for V1, HYVV-DJ for V2, C1, C2 and C3, and HYVMV-[FK1], -[OT1] and -[OT2] for C4 (Table 1). HYVV-DJ shared the highest aa sequence identities with HYVV-Jeju and -[Fukui] for V1, HYVV-Jeju for V2, C2 and C3, HYVV-[Masuda] for C1, and HYVMV-[Nar2:06] for C4 (Table 2). In an earlier study, ORFs C1 and C4 were less conserved than V1, V2, C2 and C3 between some TYLCV isolates (Navas-Castillo et al. 2000). In the case of the HYVV isolates compared with numerous begomoviruses, the difference was most pronounced in the C2, C3 and C4 regions.

To analyze the phylogenetic relationship of HYVV-Jeju and -DJ to other begomoviruses, the nt sequences of full length genome of 31 begomoviruses listed in Table 1 were aligned with those of HYVV-Jeju and -DJ using Clustal X (Jeanmougin et al. 1998). A Maximum-likelihood phylogenetic tree was drawn using PAUP*4.0b10 (Swofford 2002) with heuristic search algorithm and TbLCThV as the outgroup. Assumed nucleotide frequencies (empirical frequencies) were A=0.26286, C=0.20277, G=0.22932, and T=0.30505. Since the Hasegawa, Kishino and Yano (HKY85) model (Hasegawa et al. 1985) distinguishes between transitions and transversions (α/β) and allows unequal base frequencies, we used HKY85 model as nucleotide substitution model and performed analysis with assumed proportion of invariable sites = none and distribution of rates at variable sites = equal settings. Number of distinct data patterns under this model = 1412. The molecular clock was not enforced. Starting branch lengths were obtained using Rogers-Swofford approximation method. Trees with approximate likelihoods 5% or further from the target score were rejected without additional iteration. Bootstrapping for 1,000 replication was performed by stepwise sequence addition. On the basis of the maximum likelihood tree rooted with TbLCThV, the isolates listed in Table 1 including HYVV-Jeju and -DJ can be grouped into two major clades (Fig. 2). Clade II represents only EpYVV isolates, such as EpYVV-[MNS2], -[Yam], -[Kagawa], -[SOJ3] and -[Suya]. On the other hand, clade I is separated into five subclades A, B, C. D, and E. Subclade A forms a cluster mainly with HYMMV isolates. Subclade B is grouped with mostly HYVV isolates. Subclade C represents HYVMV-[TobkG5] and -[Miyazaki]. Subclade D contains mostly TbLCJV isolates such as TbLCJV- 1999, -2007, -[JP-2], -[JP-3] and -[Iba:06], as implied by a bootstrap value of 99%. Interestingly, subclade E contains mixed species including TbLCJV-[Koc], HYVMV-[HY12] and HYVV-[Masuda], -Jeju and -DJ. In a previous study (Ueda et al. 2008), TbLCJV, HYVMV, and EpYVV isolates in Japan formed a Far East Asian clade which was separated into two distinct subclades, TbLCJV/HYVMV and EpYVV. This separation agreed with our results of Clade I and II separation. Thus, it is clear that the two Korean isolates HYVV-Jeju and - DJ belong to Far East Asian begomoviruses. According to Ueda et al. (2008),

Table 2 Percentage nucleotide (nt) and deduced amino acid (aa) sequence identities between HYVV-DJ and other begomovirus isolates

Virus isolates[a]	Total nt	IR nt	V1 nt/aa	V2 nt/aa	C1 nt/aa	C2 nt/aa	C3 nt/aa	C4 nt/aa
HYVV-[Masuda]	89.4	69.3	91.9/96.5	90.9/91.4	92.9/94.8	91.7/89.6	90.9/87.3	94.2/86.6
TbLCJV-[Koc]	88.8	69.3	93.0/97.3	93.6/96.5	90.7/92.3	91.2/88.9	91.6/89.6	92.9/84.5
HYVMV-[HY12]	86.9	62.8	89.8/93.8	89.5/87.1	90.4/93.9	91.9/88.9	89.1/86.6	91.5/79.4
TbLJV-[Iba:06]	87.5	68.6	93.7/97.7	88.3/91.4	88.3/90.6	89.5/85.9	88.4/85.8	90.5/81.4
TbLCJV-[JP3]	87.0	67.0	92.1/96.9	89.8/87.9	89.0/92.3	89.0/84.4	88.2/83.6	91.5/78.4
HYVMV-[FK1]	86.6	65.2	85.0/86.0	87.5/90.5	92.4/93.9	89.0/86.7	88.6/84.3	93.5/86.6
TbLCJV-2007[b]	87.3	68.7	93.2/97.3	90.6/89.7	88.4/90.4	89.5/85.9	87.7/84.3	90.8/82.5
TbLCJV-1999[b]	89.1	84.8	91.1/94.9	89.5/89.7	89.2/92.8	89.2/84.4	87.4/82.8	90.5/76.3
HYVMV-[OT2]	86.7	65.1	85.1/86.8	89.5/90.5	92.5/94.5	89.0/85.9	88.9/84.3	93.5/86.6
TbLCJV-[JP2]	86.8	66.7	91.7/94.9	90.1/88.8	89.0/92.0	89.0/84.4	87.4/82.1	91.5/78.4
HYVMV-[OT1]	86.2	65.5	85.0/86.0	86.9/90.5	91.9/93.7	88.2/83.7	87.7/82.8	93.9/86.6
HYVV-[HYVVNZ1]	88.0	77.0	90.4/94.6	87.8/89.7	89.0/89.3	91.7/88.9	89.6/87.3	90.5/79.4
HYVV-[Fukui]	87.9	75.6	91.1/98.1	89.8/87.1	89.0/82.9	91.4/88.2	89.4/85.8	90.1/78.4
HYVMV-[NarI:06]	85.9	66.9	85.4/87.9	89.2/87.9	89.4/92.8	90.4/85.9	89.4/85.8	92.2/79.8
HYVMV–[TobKG5]	85.4	66.6	85.5/85.2	85.5/87.1	88.6/90.6	86.8/84.4	89.1/85.1	90.8/79.4
HYVMV-[Yam]	87.5	73.6	85.9/86.8	86.9/89.7	90.8/92.6	90.7/89.7	90.1/86.6	92.5/84.5
HYVMV-[Miyazaki]	86.5	72.8	86.1/88.7	87.8/88.8	89.9/90.6	87.3/83.7	87.9/82.1	64.8/59.3
HYVMV-[Nar2:06]	87.4	83.1	85.3/87.9	85.8/86.2	89.7/93.4	90.7/86.7	88.2/83.6	90.5/96.3
HYVMV-[SP1:00]	85.9	75.7	85.0/87.6	87.1/81.9	88.3/89.0	90.7/85.9	88.0/76.5	89.5/76.3
EpYVV-[Suya]	83.7	75.5	85.5/88.7	86.0/84.5	85.1/87.6	80.3/69.1	83.5/78.4	89.8/78.3
EpYVV-[Kagawa]	83.5	72.2	85.4/88.7	84.1/83.6	85.6/88.1	81.9/71.1	84.7/79.1	90.1/78.4
EpYVV-[Yam]	81.8	57.4	85.9/88.7	84.3/82.8	84.8/87.1	81.6/71.8	84.7/80.6	62.2/53.6
EpYVV-[MNS2]	81.3	57.9	84.5/89.1	82.6/. 84.5	84.8/87.1	81.1/71.1	84.7/78.4	62.2/53.6
EpYVV-[SOJ3]	79.0	60.7	84.8/88.7	83.5/87.1	78.2/81.5	82.4/72.6	84.4/79.1	66.3/41.2
HYVV-[UK1]	86.7	55.0	91.0/95.8	87.5/88.8	82.3/82.7	91.2/87.4	88.6/85.8	47.3/40.2
HYVV-[UK2]	87.7	76.6	90.1/94.2	87.2/87.9	89.0/89.3	91.7/88.9	89.1/86.6	90.1/78.4
SgYVGdV	77.5	64.5	74.3/79.7	72.2/65.0	82.7/84.9	75.7/61.0	80.3/74.6	84.0/67.0
PaLCuCNV	75.8	54.6	76.2/78.6	78.7/73.3	82.7/84.9	72.1/60.7	74.4/67.9	86.1/71.1
TbLCZWV	70.8	58.5	69.9/72.5	69.2/62.1	74.9/74.7	68.1/51.9	72.6/67.9	68.7/47.4
TbLCYV-[Y136]	74.0	55.8	75.2/74.7	76.3/72.3	79.7/78.6	71.6/60.0	72.6/66.4	75.2/51.6
TbLCThV	74.2	50.5	75.5/75.1	73.3/74.0	80.8/83.5	71.8/59.3	72.8/69.4	83.7/67.0
HYVV-Jeju	91.9	70.2	95.2/98.1	98.3/97.4	92.6/94.5	95.8/94.8	97.3/97.0	93.9/85.6

HYVMV-[HY12] is a possible recombinant of HYVV-[UK1] and TbLCJV-[JP3], and HYVV-[Masuda] is a possible recombinant of HYVV-[Fukui] and TbLCJV-[JP3]. Since the two Korean isolates HYVV-Jeju and -DJ clustered with these two Japanese viral isolates in the same clade, it is worth considering whether there is any recombination component in the genome of Korean isolates which might be derived from certain part(s) of DNA-A genome of other Far East Asian begomoviruses.

To further support the phygenetic relationship of HYVV-Jeju and -DJ to other begomoviruses, we also analyzed the nt sequences aligned and adjusted for PAUP analysis using MrBayes method (Ronquist and Huelsenbeck 2003). Generalised time-reversible (GTR) model of DNA substitution (Rodríguez et al. 1990) with a gamma distribution to account for rate variation among sites (Yang 1994) was used as input. Because we had no knowledge of the priors, uninformative ones were associated with branch

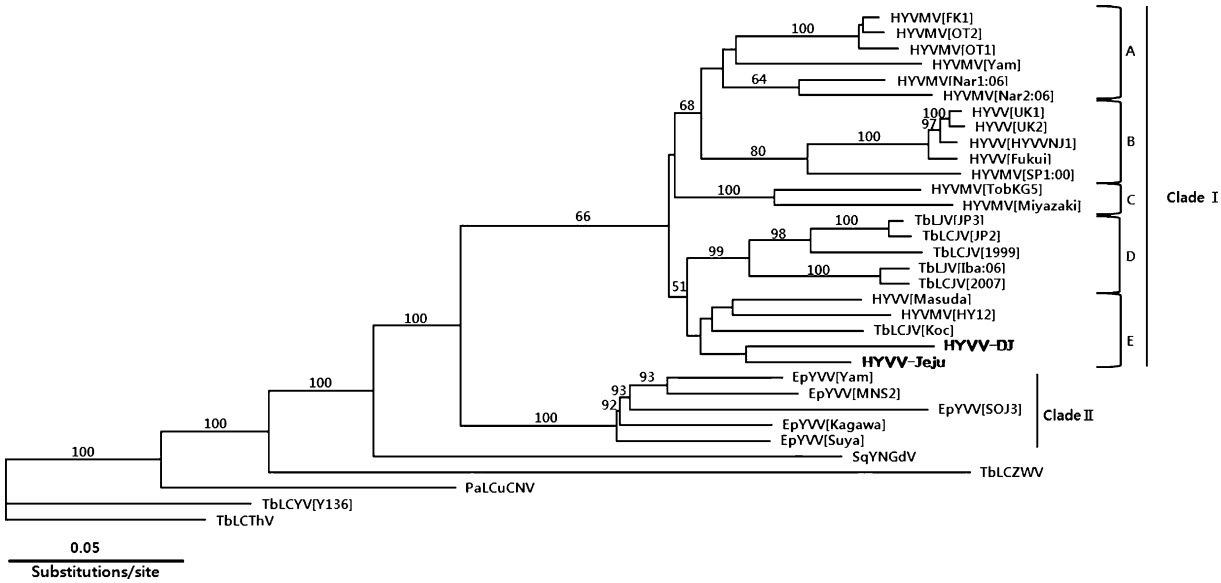

Fig. 2 Phylogenetic tree constructed based on nucleotide sequences of the full genome in begomoviruses worldwide using the maximum likelihood method in PAUP*4.0b10. The numbers at each branch indicate the percentage of 1,000 bootstraps, which support the grouping at that node. Bootstrap scores >50% are shown at major nodes. TbLCThV is the outgroup

lengths and substitution parameters of rate matrix and all trees were considered equally probable a priori (Ronquist and Huelsenbeck 2003). The Monte Carlo Markov chain algorithm in MrBayes was used to get 1,000 samples from the posterior probability distribution. A consensus tree was built from 35,000 trees with TbLCThV as the outgroup. During the run of parameters of the GTR model and the shape of gamma distribution were estimated by MrBayes with respective 95% confidence intervals. Nucleotide frequencies were $f_A = 0.252230$, $f_C = 0.194959$, $f_G = 0.235735$ and $f_T = 0.317075$; the parameters of the rate matrix were R(A–C) = 0.087254, R(A–G) = 0.251441, R(A–T) = 0.125064, R(C–G) = 0.110548, R(C–T) = 0.308709 and R(G–T) = 0.116983, and the shape parameter of the gamma distribution was alpha = 0.798501. The phylogeny of the analyzed begomoviruses by MrBayes is given in Fig. 3. The Far East Asian begomoviruses were divided into two major clades as in Fig. 2. Clade I in the phylogeny included five subclades. The Sublade A contained only HYVMV isolates and Subclade B represented only TbLCJV isolates. Subclade C has only HYVV isolates. Subclade D contains three different viruses. This is particularly interesting. As in the subclade D in Fig. 2, these three different viruses were grouped together and more closely neighbored to the HYVV isolates subclade.

The Korean HYVV-Jeju and -DJ are separated from other viruses and independently locate in Subclade E. Clade II contains the EpYVV isolates only. The grouped viral isolates in each subclade in Fig. 3 differed little from those in each subclade in Fig. 2. When we compared the phylogenetic relationship resolved in Fig. 2 and Fig. 3, the species relationship in the Far East Asian begomoviruses was better resolved by the analysis using MrBayes method with higher supporting values for tree branching.

In a previous study, Ueda et al. (2008) used neighbour joining (NJ) and the maximum-parsimony (MP) algorithms of Mega2.1 program for the phylogentic relationship analysis of the Far East Asian begomoviruses. In their work, they put HYVMV/TbLCJV isolates as a subclade and mentioned the viral subclade contained a species complex. Although they mentioned there were clusters in the HYVMV/TbLCJV subclade, they did not suggest any possible split of the species within the subclade. We found the status of subclade D in Fig. 3 is ambiguous at this moment. However, when we did pair-wise comparison of the nt sequence of TbLCJV-[Koc],

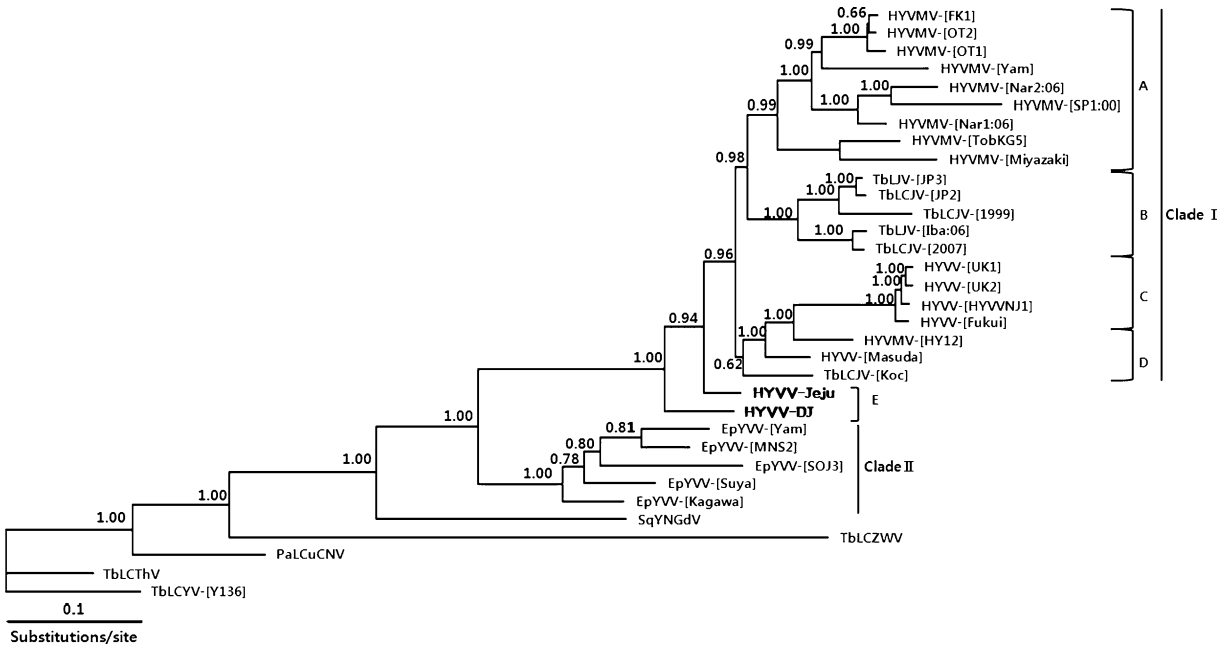

Fig. 3 The full genome of begomoviruses phylogeny inferred by Bayesian analysis corresponding to the consensus of 35,000 trees. The numbers at each branch indicate the posterior probability of each split. TbLCThV is the outgroup

HYVMV-[HY12] and HYVV-[Masuda], both HYVMV-[HY12] and TbLCV-[Koc] shared 92% sequence similarity with HYVV-[Masuda]. If we apply this 92% similarity to the geminivirus nomenclature guideline proposed by Fauquet et al. (2008), HYVMV-[HY12] and TbLCV-[Koc] will be HYVV isolates. In this case, the clade D in Fig. 3 will contain only HYVV isolates and subsequently, it will join together with the neighbouring clade C to form a clade that represents HYVV. If the clade C and D is combined as one clade, then the species complex of HYVMV, TbLCJV, and HYVV is likely

Table 3 Identification of potential recombination region of HYVV-Jeju and -DJ genomes and their parental viruses using RDP 3.44

Recombinant	Region (nt)	Parent isolates		Detected by	Average *P*-value
		Major	Minor		
HYVV-Jeju/HYVV-DJ	1129–1226	HYVMV-[SP1:00]	EPYVV-[Suya]	RDP	3.050×10^{-1}
				GENECONV	4.169×10^{-2}
				BootScan	2.500×10^{-2}
				MaxChi	–
				Chimaera	3.861×10^{-2}
				SiScan	2.177×10^{-14}
				3Seq	–
HYVV-DJ	2567–2732	HYVV-Jeju	EPYVV-[Suya]	RDP	1.157×10^{-20}
				GENECONV	1.613×10^{-18}
				BootScan	9.739×10^{-17}
				MaxChi	2.638×10^{-10}
				Chimaera	3.581×10^{-11}
				SiScan	3.835×10^{-10}
				3Seq	1.107×10^{-11}

very well sorted as each unique species. Thus, it would be necessary to revisit the relationship of HYVV-[Masuda] with HYVMV-[HY12] and TbLCV-[Koc].

Recombination is a common phenomenon among begomoviruses, contributing to the genetic diversification of the geminivirus population (Padidam et al. 1999) and playing a role in the emergence of new viruses and diseases (Varsani et al. 2008). An interspecies recombination is a predominant feature of begomovirus evolution. For example, TbLCJV (AB028604) is an interspecies recombinant in the IR between TbLCJV-[JP2] or -[JP3] and HYVMV, and TbLCJV-[Yam] has a hybrid genome in the C1 region from the major parent HYVMV and minor parent TbLCJV-[Koc] (Kitamura et al. 2004). In addition, EpYVV-[SOJ3] has a chimeric genome which may have arisen by recombination between HYVMV-[TobKG5] and EpYVV-[Yam] (Ueda et al. 2008). In earlier studies, recombination events between begomoviruses occurred in the complementary sense ORFs or IR, because both regions are more liable to form a recombination breakpoint (Kitamura et al. 2004; García-Andrés et al. 2007; Lefeuvre et al. 2007; Ueda et al. 2008). We assume that the occurrence of mixed infection and the simultaneous replication of those viruses in a same cell may have allowed for the interspecies recombination. In Japan, the high genetic diversity of TbLCV was considered to originate through the coevolution of strains, mainly within *Eupatorium* (Ooi et al. 1997). Such perennial wild host plants support mixed infection for long periods, providing more chances of the adaptation of recombinant variants. It suggested that the successful fitness of recombinants in nature may allow them to compete successfully within the population and eventually be transmitted to other hosts by the insect vector.

To investigate the possible origin of this virus and search for recombination events, the Recombination Detection Program, RDP version 3.44 (Martin et al. 2005), was used with defaulted settings (a Bonferroni corrected, high acceptable P value of 0.05, and a window size of 10) for the different detection methods. Therefore, we aligned the full genome sequence of HYVV-Jeju and -DJ with numerous begomovirus isolates after phylogenetic analysis, and identified a tract of sequence that had a potentially recombinant origin using the RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq methods implemented in RDP3.44. In order to avoid false positives, only results predicted by more than five programs including RDP (phylogeny based method) and GENECONV (substitution based method) were accepted for further analysis. In Table 3, five programs recognized the source of the recombination fragment spanning the C3 ORF (nt 1129–1226) to be close to HYVV-Jeju and -DJ, respectively. HYVV-Jeju and -DJ could have originated from the interspecies recombination between HYVMV-[SP1:00] as a major parent and EpYVV-[Suya] as a minor parent. In addition, seven programs recognized the source of another recombination fragment spanning the IR (nt 2567–2732) to be close to HYVV-DJ. Interestingly, the analysis indicated that HYVV-DJ is a recombinant originating from HYVV-Jeju as a major parent and EpYVV-[Suya] as a minor parent. Therefore, HYVV-DJ is a recombinant virus combining parts of the genomes of HYVMV-[SP1:00], HYVV-Jeju and EpYVV-[Suya]. It suggested that HYVV was established in one area after recent adaptation and spread to the other area after rapid variation. At present, it is not clear how theses HYVV isolates have been transmitted from Japan to Korea. Considering that the climate of Korea has been gradually changing from temperate to subtropical climate and the whitefly vector also has been spreading not only to tomato but also to other *Solanaceae* crops across the country, it is expected that new variants of HYVV would be continuously generated. Thus, it will be necessary to further investigate the occurrence of HYVV in different host plants and at different geographical sites in Korea.

Acknowledgements This study was supported by the Research Center for Tomato Export (Sangmyung University, Cheonan) and by the Technology Development Program for Agriculture and Forestry in the Ministry for Food, Agriculture, Forestry and Fisheries, Republic of Korea

References

Fauquet, C. M., & Stanley, J. (2005). Revising the way we conceive and name viruses below the species level: a review of geminivirus taxonomy calls for new standardized isolate descriptors. *Archives of Virology, 150*, 2151–2179.

Fauquet, C. M., Briddon, R. W., Brown, J. K., Moriones, E., Stanley, J., Zerbini, M., et al. (2008). Geminivirus strain demarcation and nomenclature. *Archives of Virology, 153*, 783–821.

Froger, A., & Hall, J. E. (2007). Transformation of plasmid DNA into *E. coli* using the heat shock method. *Journal of Visualized Experiments, 6*, 253.

García-Andrés, S., Tomas, D. M., Sanchez-Campos, S., Navas-Castillo, J., & Moriones, E. (2007). Frequent occurrence of recombinants in mixed infections of tomato yellow leaf curl disease-associated begomoviruses. *Virology, 359*, 302–312.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., & Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences, 23*, 403–405.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., & Robertson, D. (2000). Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Biochemistry and Molecular Biology, 35*, 105–140.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution, 22*, 160–174.

Kitamura, K., Murayama, A., & Ikegami, M. (2004). Evidence for recombination among isolates of *Tobacco leaf curl Japan virus* and *Honeysuckle yellow vein mosaic virus*. *Archives of Virology, 149*, 1221–1229.

Lazarowitz, S. G. (1992). Geminiviruses: genome structure and gene function. *Critical Reviews in Plant Sciences, 11*, 327–349.

Lefeuvre, P., Martin, D. P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., et al. (2007). Begomovirus 'melting pot' in the south-west Indian Ocean islands: molecular diversity and evolution through recombination. *The Journal of General Virology, 88*, 3458–3468.

Navas-Castillo, J., Sánchez-Campos, S., Noris, E., Louro, D., Accotto, G. P., & Moriones, E. (2000). Natural recombination between Tomato yellow leaf curl virus-Is and Tomato leaf curl virus. *The Journal of General Virology, 81*, 2797–2801.

Martin, D. P., Williamson, C., & Posada, D. (2005). RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics, 21*, 260–262.

Ogawa, T., Sharma, P., & Ikegami, M. (2007). First report of a strain of Tobacco leaf curl Japan virus associated with a satellite DNA in honeysuckle in Japan. *New Disease Reports* http://www.bspp.org.uk/ndr/July 2007/2007-14.asp.

Ogawa, T., Sharma, P., & Ikegami, M. (2008). The begomoviruses Honeysuckle yellow vein mosaic virus and Tobacco leaf curl Japan virus with DNAβ satellites cause yellow dwarf disease of tomato. *Virus Research, 137*, 235–244.

Ooi, K., Ohshita, S., Ishii, I., & Yahara, T. (1997). Molecular phylogeny of geminivirus infecting wild plants in Japan. *Journal of Plant Research, 110*, 247–257.

Padidam, M., Sawyer, S., & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology, 265*, 218–225.

Pietersen, G., Idris, A. M., Krüger, K., & Brown, J. K. (2008). Characterization of *Tomato curly stunt virus*: a new tomato-infecting begomovirus from South Africa. *Plant Pathology, 57*, 809–818.

Rodríguez, F., Oliver, J. L., Marín, A., & Medina, J. R. (1990). The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology, 142*, 485–501.

Ronquist, F., & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*, 1572–1574.

Saunders, K., Bedford, I. D., Briddon, R. W., Markham, P. G., Wong, S. M., & Stanley, J. (2000). A unique virus complex causes Ageratum yellow vein disease. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 6890–6895.

Shimizu, S., & Ikegami, M. (1999). Complete nucleotide sequence and the genome organization of tobacco leaf curl geminivirus from Japan. *Microbiology and Immunology, 43*, 989–992.

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version Version 4.0b10*. Sinauer Associates, Sunderland, Mass.

Ueda, S., Kimura, T., Onuki, M., Hanada, K., & Iwanami, T. (2004). Three distinct groups of isolates of Tomato yellow leaf curl virus in Japan and construction of an infectious clone. *Journal of General Plant Pathology, 70*, 232–238.

Ueda, S., Onuki, M., Hanada, K., & Takanami, Y. (2008). Unique grouping of the Far East Asian begomovirus complex based on sequence analysis of the DNA-A genome and associated DNAβ satellite molecules isolated from tomato, honeysuckle and *Eupatorium* plants in Japan. *Archives of Virology, 153*, 417–426.

Varsani, A., Shepherd, D. N., Monjane, A. L., Owor, B. E., Erdmann, E. P., Rybicki, E. P., et al. (2008). Recombination, decreased host specificity and increased mobility may have driven the emergence of maize streak virus as an agricultural pathogen. *The Journal of General Virology, 89*, 2063–2074.

Were, H. K., Takeshita, M., Furuya, N., & Takanami, Y. (2005). Molecular characterization of new begomovirus infecting honeysuckle in Kobe. *Journal of the Faculty of Agriculture Kyushu University, 50*, 61–71.

Yahara, T., Watanabe, K., & Kawahara, T. (1995). Eupatorium. In K. Iwatsuki, T. Yamazaki, D. E. Boufford, & H. Ohba (Eds.), *Flora of Japan, vol IIIb* (pp. 110–114). Tokyo: Kodansha Scientific.

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution, 39*, 306–314.