Mol. Cells *30*, 467-476, November 30, 2010
DOI/10.1007/s10059-010-0143-7

Molecules and Cells

Phylogenetic Analysis and Inflow Route of *Tomato Yellow Leaf Curl Virus* (TYLCV) and *Bemisia tabaci* in Korea

Hyejung Lee[1,6], Woogeun Song[2,6], Hae-Ryun Kwak[3], Jae-deok Kim[1,3], Jungan Park[1], Chung-Kyoon Auh[4], Dae-Hyun Kim[5], Kyeong-yeoll Lee[2], Sukchan Lee[1,*], and Hong-Soo Choi[3,*]

Tomato yellow leaf curl virus **(TYLCV) is a member of the genus *Begomovirus* of the family *Geminiviridae*, members of which are characterized by closed circular single-stranded DNA genomes of 2.7-2.8 kb in length, and include viruses transmitted by the *Bemisia tabaci* whitefly. No reports of TYLCV in Korea are available prior to 2008, after which TYLCV spread rapidly to most regions of the southern Korean peninsula (Gyeongsang-Do, Jeolla-Do and Jeju-Do). Fifty full sequences of TYLCV were analyzed in this study, and the AC1, AV1, IR, and full sequences were analyzed via the muscle program and bayesian analysis. Phylogenetic analysis demonstrated that the Korea TYLCVs were divided into two subgroups. The TYLCV Korea 1 group (Masan) originated from TYLCV Japan (Miyazaki) and the TYLCV Korea 2 group (Jeju/Jeonju) from TYLCV Japan (Tosa/Haruno). A *B. tabaci* phylogenetic tree was constructed with 16S rRNA and mitochondria cytochrome oxidase I (MtCOI) sequences using the muscle program and MEGA 4.0 in the neighbor-joining algorithm. The sequence data of 16S rRNA revealed that Korea *B. tabaci* was closely aligned to *B. tabaci* isolated in Iran and Nigeria. The Q type of *B. tabaci,* which was originally identified as a viruliferous insect in 2008, was initially isolated in Korea as a non-viruliferous insect in 2005. Therefore, we suggest that two TYLCV Japan isolates were introduced to Korea via different routes, and then transmitted by native *B. tabaci.***

INTRODUCTION

Members of the *Geminiviridae* family evidence circular single-stranded DNA (ssDNA) genomes, and are divided into four genera (*Mastrevirus, Curtovirus, Begomovirus, and Topocuvirus*) based on genome composition, insect vector, and host range (Zhang et al., 2009). The *Tomato yellow leaf curl virus* (TYLCV) is a member of the genus *Begomovirus* and has an approximately 2.7-2.8 kb genome, with monopartite or bipartite genome compositions. The monopartite virus has a DNA-A genome harboring six open reading frames (ORFs) that encode for regulation-related genes (AC1-AC4) and structure genes (AV1 and AV2). The bipartite viruses consist of two components, DNA-A and DNA-B (Kheyr-Pour et al., 1991). The DNA-A genome encodes for four proteins associated with the regulation of replication, encapsidation, overcoming host defense, and the control of gene expression (AC1-AC4) (Briddon et al., 2010), whereas DNA-B encodes for two proteins (BV1 and BV2) associated with intra- and intercellular movement in host plants (Briddon et al., 2010).

TYLCV is the primary begomovirus infecting tomatoes, and causes 100% losses in tomato crops grown in tropical and subtropical regions (Yongping et al., 2008). TYLCV-infected tomatoes evidence typical disease symptoms including leaf curling, yellowing, and stunted growth. After being isolated and characterized, TYLCV has evidenced a rapid global spread, and TYLCV diseases have been reported in Japan, Korea, China, America, Israel, Egypt, and Australia, are currently and consistently spreading to new regions (Yongping et al., 2008; Zhang et al., 2009). Briddon et al. (2010) reported that monopartite geminiviruses originated from the Old World (Africa, India, Asia, and Japan) and bipartite geminiviruses originated from the New World (Latin-America and Meso-America). Generally, bipartite geminiviruses evolved from monopartite geminiviruses and the ORFs of the monopartite genome were split into two genomes or one genome and a DNA β-satellite. TYLCV has a monopartite genome, and studies involving the phylogenetic and evolutionary analysis of TYLCV have been conducted in a variety of different countries, including Japan (Shimizu and Ikegami, 1999), Iran (Fazeli et al., 2009), China (Zhang et al., 2009), and Thailand (Sawangjit et al., 2005), on the basis of the TYLCV sequences isolated from virus-infected plants. Phylogenetic analysis based on virus sequence analyses among *begomovirus* groups have been conducted previously in *Sweet potato leaf curl virus* (SPLCV) (Luan et al., 2007),

[1]Department of Genetic Engineering, Sungkyunkwan University, Suwon 440-746, Korea, [2]School of Applied Biosciences, Kyungpook National University, Daegu 702-701, Korea, [3]Agricultural Microbiology Division, National Academy of Agricultural Science, Suwon 441-707, Korea, [4]Department of Biological Sciences, Mokpo National University, Muan 534-729, Korea, [5]Fruit Research Division, National Institute of Horticultural and Herbal Science, Suwon 440-706, Korea, [6]These authors contributed equally to this work.
*Correspondence: cell4u@skku.edu (SL); hschoi@korea.kr (HSC)

Received June 30, 2010; revised August 16, 2010; accepted August 26, 2010; published online October 14, 2010

Keywords: begomovirus, *Bemisia tabaci,* phylogeny, tomato, TYLCV

Cotton leaf curl virus (CLCV) (Shahid et al., 2007), and *Tomato leaf curl virus* (ToLCV) (Matsuda et al., 2008). TYLCV Iran isolates were analyzed to determine the phylogeny of TYLCV on the basis of conserved sequences in AC1 (Rep), AV1 (CP), and also on full sequences (Fazeli et al., 2009; Sawangjit et al., 2005).

TYLCV was initially isolated in Tongyeong, Gyeongsang-Do (the southern part of the Korean peninsula) in 2008 and spread rapidly to the entirety of the Gyeongsang-Do, Jeolla-Do, and Jeju-Do regions. TYLCV is transmitted by the *B. tabaci* whitefly (Ghanim et al., 1998; Navot et al., 1991). More than 20 biotypes of *B. tabaci* have been identified worldwide, and the classification and phylogenetic analysis of *B. tabaci* were conducted via analyses of 16S ribosomal RNA (Everett et al., 2005; Thao et al., 2003), mitochondrial cytochome oxidase I (MtCOI) (Boykin et al., 2007; Hsieh et al., 2007; Kim and Lee, 2008; Shatters et al., 2009; Tsagkarakou et al., 2007), ribosomal ITS1 (De Barro et al., 2000), and microsatellites (Tsagkarakou et al., 2007). These sequences are, in most species, sufficiently conserved to trace the phylogenetic relationship among the biological species, but evidence sufficient diversity in specific regions for phylogenic comparisons (Hillis and Dixon, 1991). Therefore, genetic diversity and identification of species origins can be conducted via the analysis of sequence variations within conserved sequences (Tsagkarakou et al., 2007).

In this study, we isolated and analyzed 50 full sequences of TYLCV genomes from TYLCV-infected tomatoes, and compared the sequence variations among viral ORFs (AC1 and AV1) and intergenic regions (IR) with those of known TYLCV sequences isolated from Asian countries including China, Japan, Thailand, and Indonesia. Additionally, sequences of 16S rRNA and MtCOI isolated from *B. tabaci*, an insect vector of TYLCV, were analyzed. The inflow routes of TYLCV and *B. tabaci* were reviewed by analyzing the phylogeny of TYLCV and *B. tabaci* isolated and collected in Korea.

MATERIALS AND METHODS

Collection of TYLCV-infected tomatoes and *Bemisia tabaci*

TYLCV-infected tomatoes *(Lycopersicon esculentum)* exhibiting leaf curling, yellowing, and stunted growth were collected from more than 25 tomato fields in Korea, including the Gyeongsangnam-Do, Jeolla-Do, and Jeju-Do regions in 2008-2009 (Fig. 1B). *B. tabaci* were also collected from approximately 20 tomato farms within Gyeongsang-Do, Jeolla-Do, and Jeju-Do.

Genomic DNA isolation

Total genomic DNA of tomatoes and *B. tabaci* were extracted via the method developed by (Dellaporta et al., 1983) or a modified method. Genomic DNA extraction was conducted after the tissues were ground using Kontes pestles (K749541-0000, Thermo Fisher Scientific Co., USA) in Eppendorf tubes. After washing in 70% ethanol, the DNA pellets were dried at room temperature for 10 min. Finally, the pellets were resuspended in approximately 50 μl of distilled water. After determining the concentration and purity of the DNA, the extracted total DNA was stored at -80°C.

Southern hybridization

Equal amounts of genomic DNA were loaded onto 1.0% agarose gel and electrophoretically separated for 6 h at 60 V. Depurination, denaturation, and neutralization were conducted according to the protocol described by (Lee et al., 1994). DNA was transferred to a nylon membrane and incubated with prehybridization solution (Rapid-hyb buffer, Amersham Biosciences, NIF939) containing salmon sperm DNA (100 ng/ml) for 2 hrs at 65°C. ^{32}P-dCTP labeled probes were prepared from the TYLCV CP gene using the Amersham Red prime II Random primer labeling system (GE healthcare, RPN1633) and hybridization was conducted for 16 h at 65°C. The nylon membranes were then consecutively washed with 50 ml 2× SSC and 0.1% SDS followed by 1× SSC and 0.1% SDS for 1 h each. The nylon membrane was exposed to X-ray film and maintained in a deep freezer for 1-2 days.

Polymerase chain reaction (PCR)

The primers for the positional cloning of the TYLCV genome were designed from representative TYLCV genome sequences deposited with NCBI using the Primer 3 software. Using the sequence information, we designed primer pairs to amplify 0.7, 0.8, 1.0 and 0.6 kb fragments (Table 1). The reaction mixes contained 2.0 μg template DNA (genomic extracts), 5 pmoles of each forward and reverse primer (final concentration 0.25 pmoles per reaction), 10× PCR buffer, 2.5 mM dNTP (final concentration 0.125 μM), 1.5 U SUPERTAQ-Plus (*Taq* polymerase mixed with *Pfu* polymerase, SuperBio Co., Korea) and ddH_2O to yield a total reaction volume of 20 μl. The PCR conditions were as follows: 95°C for 3 min for pre-denaturation; 35 cycles of 94°C for 1 min for denaturation, 55°C for 1 min for primer annealing, 72°C for 1 min for extension; and 72°C for 10 min for final extension. The PCR products were stored at 8°C prior to analysis on 1.0% agarose gel. Sequencing of the TYLCV full genome identified six TYLCV ORFs using vector NTI 11.0 software. *B. tabaci* 16S rRNA and MtCOI were amplified with specific primer sets (Table 1) under the same PCR reaction conditions as were employed in the TYLCV study. PCR products were analyzed on 1.0% agarose gel and sequenced for phylogenetic analysis.

Sequence identity analysis of TYLCV

The identity of TYLCV sequences was analyzed using a multialign program (http://multalin.toulouse.inra.fr/multalin/multalin.html) with representative isolates from Japan (AB192965 and AB116634), China (NC_004044), Vietnam (EU189150), Thailand (NC_000869), Indonesia (NC_008267), Iran (AJ132711), Israel (AB110217), Egypt (AY597174), and Korea Masan (HM130912). The two out-groups were analyzed by Taiwan ToLCV (GU723730) and Korea *Tobacco leaf curl virus* (TbLCV) Sunchang (HM164550).

Phylogenetic analysis of TYLCV and *Bemisia tabaci*

All sequences were aligned using the muscle program. To construct a phylogeny tree, replication associated protein (Rep, AC1), coat protein (CP, AV1), intergenic region (IR), and complete TYLCV sequences from China, Japan, Thailand, Vietnam, Indonesia, and Korea were analyzed by MrBayes 3.0 and neighbor-joining to MEGA 4.0 (Tamura et al., 2007). Bayesian software MrBayes 3.0 Four Metropolises coupled with Markov Chain Monte Carlo (MCMC) chains were run, stopping when the standard divergence of the split frequencies was less than 0.01 (Kim et al., 2009). The ORFs were analyzed over 5 million generations and the full sequences over 10 million generations. In both cases, four were sampled every 100 generations and the first 25% burn in (SUMP, SUMT) cycles were discarded prior to the construction of the consensus tree. Consensus trees were visualized with MEGA 4.0 (Paredes-Esquivel et al., 2009). The 16S rRNA and MtCOI sequences of *B. tabaci* collected from Japan, China, Turkey, Pakistan, India, and USA were also analyzed by the muscle program followed by MEGA

Table 1. Information on primers specific for TYLCV and *Bemisia tabaci*

	Primer name	Sequences	Expected loci (b)	Expected size (kb)
TYLCV sequence primer	TYLCV1F	5′-GTCAACCAATCAAATTGCATCCTCAA-3′	61-83	0.71
	TYLCV1R	5′-GTCCAAAATCCATTGGGC-3′	773-756	
	TYLCV2F	5′-CGGGATAGGTTTCAAGTGATGAGGA-3′	830-854	0.87
	TYLCV2R	5′-CAAGACCTACCTCCTCATATAAGGAA-3′	1683-1702	
	TYLCV3F	5′-CTACACGCTTACGCCTTATTGG-3′	1534-1555	1.04
	TYLCV3R	5′-CCTAACATATCCCAATTGTTCTCTCTC-3′	2551-2577	
	TYLCV4F	5′-CAGGTCGAAGAACCGTTGG-3′	2387-2406	0.63
	TYLCV4R	5′-GTAAGTTCCTCAACGGCCTG-3′	243-222	
TYLCV specific primer	TYLCV_Sp AC1	5′-CGCCTTATTGGTTTCTTCTTG-3′	1545-1565	1.6
	TYLCV_Sp AV2	5′-AAACTTACGAGCCCAATACA-3′	115-134	
Bemisia tabaci 16S rRNA	LR-J	5′-CCGGTCTGAACRCAGATCA-3′	12887-12905	0.4
	LR-N	5′-GCCTGTTTACCAAAAACAT-3′	13398-13416	
Bemisia tabaci MtCOI	BeTab F2	5′-GTTGTTACTTCTCATGCTTTC-3′	163-183	1.07
	BeTab R2	5′-GACACCAGGTTATAATTGTTT-3′	1212-1232	

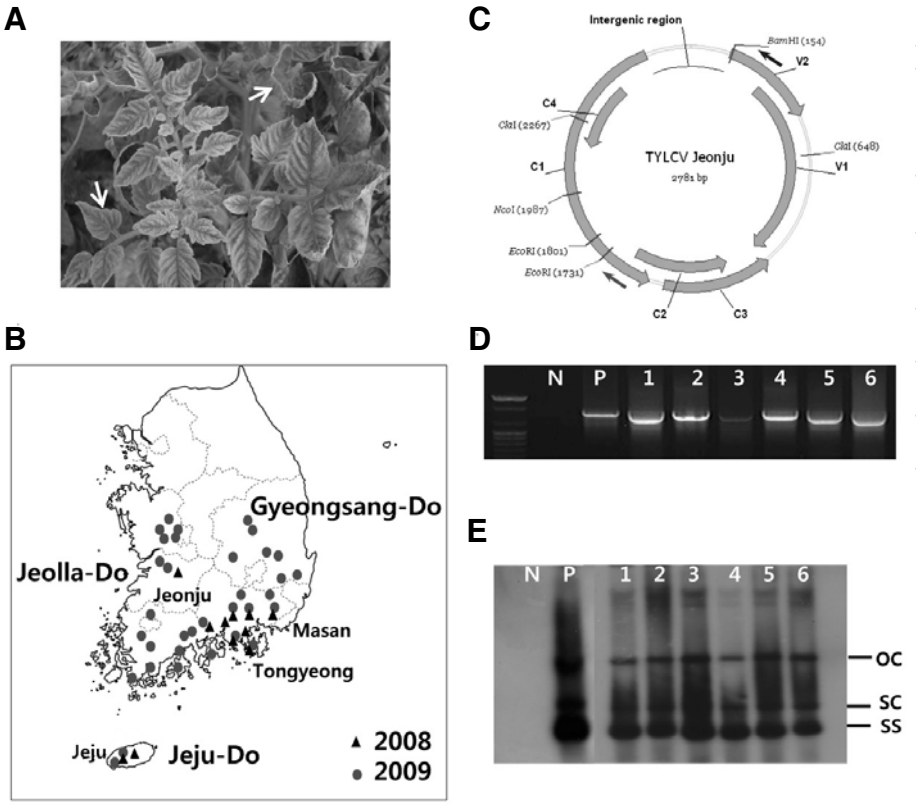

Fig. 1. Genome composition and detection of TYLCV in Korea isolates. (A) TYLCV-infected tomatoes from Korea exhibited symptoms of curling, yellowing, and stunted growth. (B) The first infected tomato was isolated from Tongyeong in 2008 and subsequently spread to other regions. TYLCV-infected tomatoes have been confirmed in Gyeongsang-Do, Jeolla-Do, and Jeju-Do. (C) Analysis of full sequences of TYLCV from infected tomatoes showing different regions of the genome and the TYLCV-specific primers designed for PCR detection. (D) PCR analysis of TYLCV-infected tomatoes (target size; 1.6 kb). (E) Southern blot analysis of TYLCV replication forms using TYLCV-specific probe. Positive (P), TYLCV Samcheonpo; lane 1, Tongyeong; lane 2, Masan; lane 3, Jeonju; lane 4, Jeju; lane 5, Haenam; lane 6, Gimhae.

4.0 programs after neighbor-joining to draw the phylogenetic trees in MEGA 4.0 (Luan et al., 2007; Tamura et al., 2007). Bootstraps of 1000, both of 0.05 present the *P*- value.

RESULTS

Complete sequence analysis of TYLCV-Korea

TYLCV-infected tomatoes evidenced typical disease symptoms, including leaf curling, yellowing, and stunted growth (Fig. 1A). TYLCV-infected tomatoes were initially identified in Tongyeong, Gyeongsang-Do, Korea in June 2008 and TYLCV subsequently spread rapidly to tomato fields in southern regions of Korea. TYLCV-infected tomatoes were reported in Jeolla-Do and Jeju-Do prior to 2009 (Fig. 1B). TYLCV from tomato tissues was collected from Tongyeong, Haman, Jeju, and Jeonju, and a total of 50 complete sequences were analyzed after PCR with primer sets designed for TYLCV sequence analysis (Table 1). ORF mapping analysis in vector NTI version 11.0 demonstrated that TYLCV has two virion-sense ORFs (AV1 and AV2) and four complementary-sense ORFs (AC1, AC2, AC3, and AC4). A -300 base intergenic region (IR) containing a hairpin structure (TAATATT↓AC) was detected between AC1 and AV2 ORF (Fig. 1C). To detect TYLCV in tissues by PCR, a TYLCV-specific primer set for PCR was designed on the basis of the

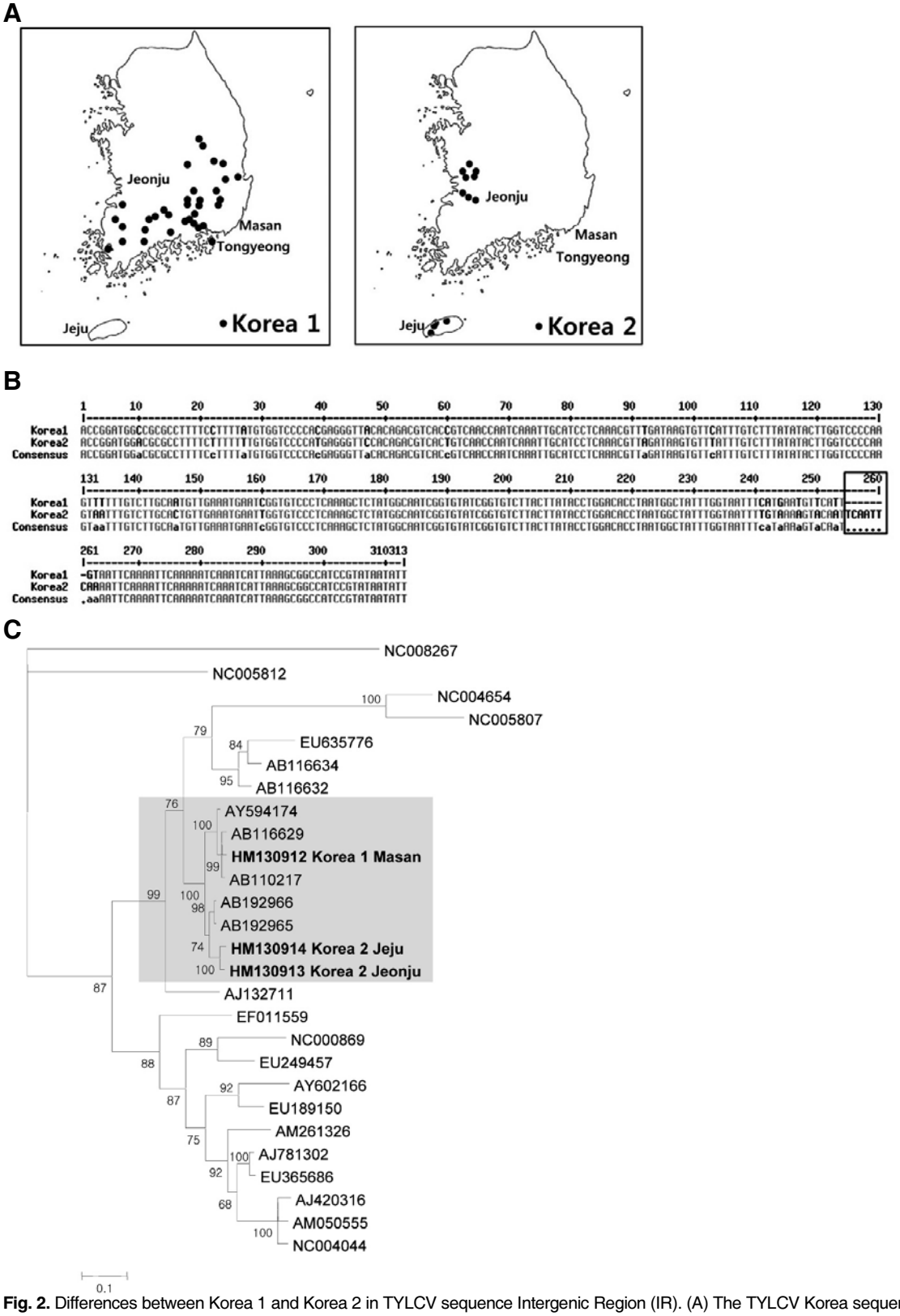

Fig. 2. Differences between Korea 1 and Korea 2 in TYLCV sequence Intergenic Region (IR). (A) The TYLCV Korea sequences identified two groups: the first group had a 2774 Nucleotide genome and was isolated from the Masan region (Gyeongsang-Do and Jeonllanam-Do) and the second group had a 2871 Nucleotide genome and was isolated from the Jeju and Jeonju regions (Jeju-Do and Jeonllabuk-Do) (B) Comparison of the IR sequences of groups 1 and 2. The Korea 1 (Masan) isolate was confirmed to have a deletion of six bases. (C) The phylogenetic tree of TYLCV IR determined for the isolates in Table 2 using muscle and bayesian analysis.

Table 2. Percentage nucleotide and amino acid sequence identity of TYLCV Jeonju (Korea 2 group) and other begomoviruses

	AC1		AC2		AC3		AC4		AV1		AV2		IR
	nt [a]	aa[b]	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa	nt
Korea Masan (HM130912)	98.2	97.7	97.5	96.2	97.2	97.0	98.9	96.9	98.1	99.2	98.2	100	91.3
Japan-Israel (AB192965)	99.0	98.3	99.0	97.7	99.2	99.2	98.9	96.9	99.3	99.6	99.0	100	96.8
Japan-Mild (AB116634)	86.5	98.3	99.0	97.7	98.7	99.2	75.5	96.9	98.7	99.6	98.8	100	96.8
Egypt (AY594174)	98.4	98.0	97.7	96.2	97.5	98.5	98.2	94.8	98.9	99.2	98.5	98.2	94.6
Israel (AB110217)	98.5	98.0	97.7	97.0	97.5	97.7	81.0	96.9	98.3	99.2	98.2	98.2	76.8
Iran (AJ132411)	85.1	89.4	97.3	97.0	96.0	94.7	80.6	68.6	97.2	98.4	97.4	97.4	76.8
China (NC_004044)	79.3	82.8	75.2	68.1	77.0	75.3	83.6	70.1	85.4	82.5	74.6	77.5	58.4
Vietnam (EU189150)	80.5	86.3	74.3	71.1	73.8	74.6	82.9	69.0	76.2	84.8	75.7	81.8	65.0
Thailand (NC_000869)	71.1	76.9	80.0	66.9	74.3	69.5	47.7	37.1	71.9	79.8	71.7	73.2	51.3
Indonesia (NC_008267)	70.1	77.2	71.5	66.6	73.5	71.6	51.2	37.1	73.1	79.8	71.7	73.2	38.9
Korea TbLCV (HM164550)	77.7	83.1	66.6	48.5	72.3	73.8	80.9	68.0	70.1	77.5	70.4	74.1	53.2
Taiwan ToLCV (GU723730)	79.7	85.9	74.2	69.2	73.0	76.8	83.6	74.2	75.5	82.9	69.8	75.0	51.4

[a]nucleotide sequence identity
[b]amino acid sequence identity

conserved sequences at AC1 and AV2, and the sequences of the IR regions were analyzed (Figs. 1C and 1D). Additionally, TYLCV replication was evaluated via southern hybridization analysis with specific probes prepared from the TYLCV coat protein gene (AV1). Southern hybridization data revealed the single-stranded DNA (ssDNA) typical of a geminivirus genome and double-stranded DNAs [super coiled (SC) dsDNA and open circular (OC) dsDNA] as products of the rolling circle replication mechanism (RCR) (Fig. 1E).

Sequence homology analysis of TYLCV and other DNA-A Begomoviruses

TYLCV in Korea was divided into two groups based on sequence and size differences. The TYLCV Korea 1 group has a genome size of 2774-nucleotide and was isolated from Gyeongsang-Do and Jeollanam-Do. The TYLCV Korea 2 group has a 2781-nucleotide genome and was isolated from Jeju-Do and Jeollabuk-Do (Fig. 2A). TYLCV Korea 1 group lacked 7 bases of DNA present in the TYLCV Korea 2 group between bases 255-261 in the IR region (Fig. 2B). Interestingly, the results of phylogeny analysis demonstrated that the TYLCV Masan isolate was separated from TYLCV Korea 1 group when the TYLCV sequences from China and Japan were compared with those of the TYLCV Korea 1 group (Fig. 2C). A phylogenetic tree was constructed using bayesian methods based on the reported TYLCV IR sequences isolated from Korea and other countries (TYLCV sequences from Japan, China, Vietnam, Thailand, Indonesia, Iran, Israel, and Egypt; ToLCV sequences from China and Taiwan) (Supplementary Table 1). Phylogenetically, the TYLCV Korea 1 group (Masan) was linked to the TYLCV from Japan Miyazaki (AB116629) and TYLCV Israel (AB110217). The TYLCV Korea 2 group (Jeju/Jeonju) was linked to TYLCV from Japan; Israel isolates (AB192966 and AB192965) (Fig. 2C). Table 2 shows the identity of the amino acid and DNA sequences of each viral ORF among viruses that were closely related geographically and evolutionally. When the sequences of all viral ORFs were compared, TYLCV Korea 1 group evidenced 2-3% sequence differences from the TYLCV Korea 2 group (Jeju/Jeonju). The amino acid sequences of all ORFs also evidenced 2-4% differences between the TYLCV Korea 1 and 2 groups, except for ORF AV2 (Table 2). The isolate closest to TYLCV Korea was TYLCV Japan; followed by the Israel (AB192965 and AB110217), and Iran (AJ132411) isolates. When the sequence identities of the TYLCV Korea 2 group and the Japan isolates (AB192965) were compared, both amino acid and nucleotide sequences shared 96-99% homology (Table 2). Another TYLCV Japan (AB116634) that induced mild symptoms in tomatoes evidenced more than a 10% difference in nucleotide sequences and an identical amino acid sequence to TYLCV Japan Israel (AB192965) (Table 2). The AC3 sequence of the TYLCV Japan mild isolate (AB116634) was 25% different from the TYLCV Korea 2 group. Other TYLCVs from China, Vietnam, and Indonesia, as shown in Table 2 were geographically distant from Japan, and thus the homology of ORF sequences among them was as low as 83%. The IR region of TYLCV Korea 2 group shared a homology of approximately 60% with isolates from China (NC_004044), Thailand (NC_000869), and Indonesia (NC_008267).

Phylogenetic relationships of TYLCV sequences

In an effort to evaluate the evolutionary relationships between viruses, phylogenetic analyses were conducted with the sequences of IR (Fig. 2C), the ORFs of AC1 (Fig. 3A) and AV1 (Fig. 3B) and the full genome (Fig. 4) from TYLCV Korea and other viruses from adjacent countries, including eight China TYLCV isolates, one Taiwan ToLCV isolate, five Japan TYLCV isolates, one Vietnam TYLCV isolate, one Indonesia TYLCV isolate, three Thailand TYLCV isolates, two Iran TYLCV isolates, one Israel TYLCV isolate, and a single Egypt TYLCV isolate. The *Honeysuckle yellow vein virus* (HYVV) (NC_005807) and *Pepper yellow vein virus* (PYVV) (AY502935) were used as out-group controls (Supplementary Table 1) using the muscle program for multi-alignment and the bayesian method for phylogeny analysis. The AC1 sequences of TYLCV were divided into two groups. The first group contained the TYLCVs of China, Thailand, and Vietnam, and the second contained TYLCV isolates from Israel, Iran, Egypt, and Japan (Fig. 3A). TYLCV Korea 1 group (Masan) was evolutionarily closely related to TYLCV Japan Miyazaki (AB116629) and

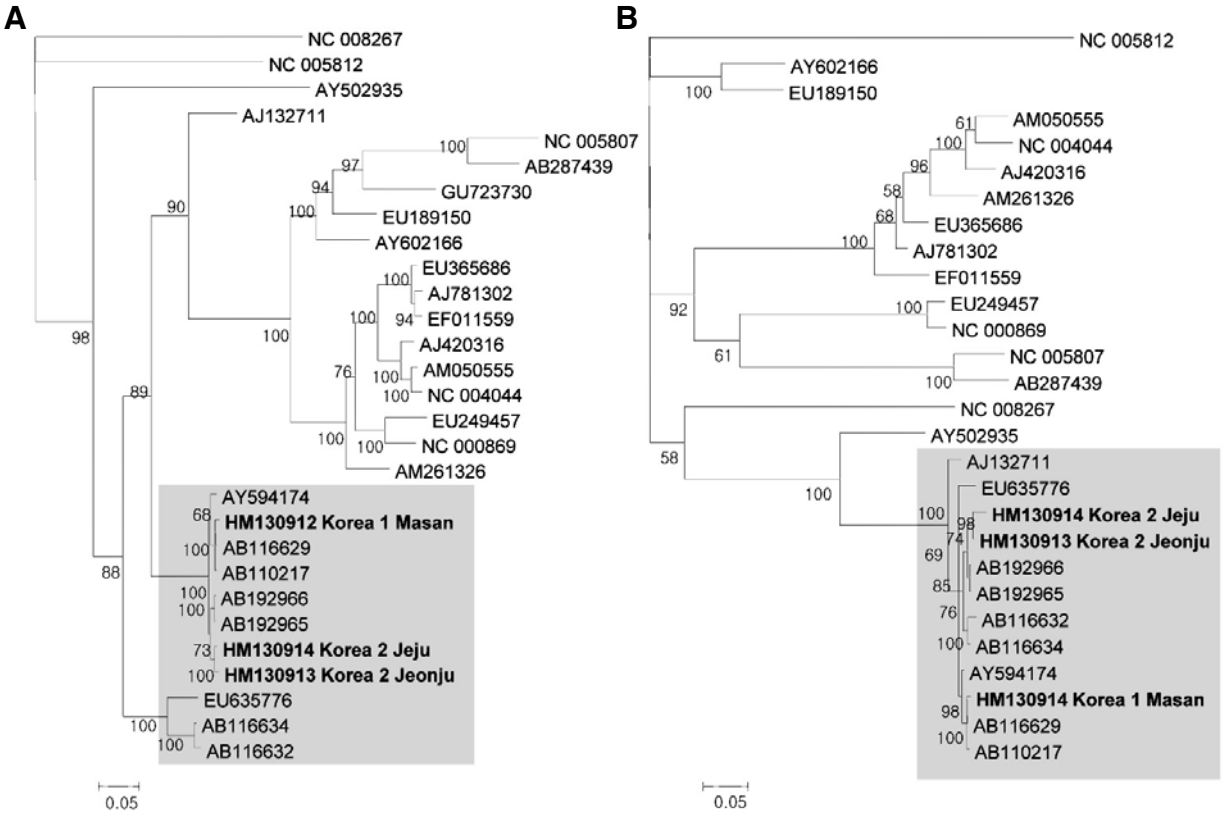

Fig. 3. Phylogenetic tree of AC1 and AV1 in TYLCV using Bayesian analysis. Phylogeny analysis used Bayesian methods for the MCMC analysis of 5 million generations and cut off 25% burn-in. The scale bar for each indicates the horizontal distance per site (0.05). The out-groups employed were the *Tomato yellow leaf curl Kanchanaburi virus* (NC_005812), and the *Tomato yellow leaf curl Indonesia virus* (NC_008267). (A) The AC1 and (B) AV1 se-quences of Korea 1 (Masan) and Korea 2 (Jeju/Jeonju) TYLCV isolates were compared with those of other countries (China, Japan, Viet-nam, and Thailand).

TYLCV Israel (AB110217) (Fig. 3A). Members of the TYLCV Korea 2 group (Jeju/Jeonju) have a 2781-nucleotide genome, and thus were positioned adjacent to one another (Figs. 3 and 4). Additionally, the TYLCV Korea 2 group was aligned closely to the TYLCV Israel isolates [AB192966 (Japan Haruno, 2005) and AB192965 (Japan Tosa, 2005)]. However, the China and Vietnam TYLCVs were phylogenetically aligned at a distance from the TYLCV Korea 1 and 2 groups (Fig. 3A). Sequence analysis of AV1 demonstrated that the TYLCVs were divided into two groups as shown in the AC1 data (Fig. 3B). The phylogeny data of the IR region, which is known to be the most variable region in geminiviruses, were identical to those of AC1and AV1. Finally, the results of a full sequence analysis of TYLCV were similar to the AC1, AV1, and IR results (Fig. 4). Geographically, Korea and Japan are quite close to one another, and the commercial traffic between the two countries grows every year. Therefore, when the TYLCV sequences of only Korea and Japan were compared, three Japan isolates [Miyazaki (AB116629), Haruno (AB192966), and Tosa (AB-192965)] were aligned closely with the TYLCV Korea isolates (Figs. 3A and 4).

Phylogenetic analysis of 16S rRNA and MtCOI in *Bemisia tabaci*

In an effort to characterize the relationship between TYLCV and its insect vector, *B. tabaci,* whitefly specimens were collected from 13 TYLCV-infected tomato fields in various regions of Korea, including Jeju-Do (Ilgwari, Deoksuri, Youngnakri, and Josuri), Gyeongsang-Do (Daegu, Seongju, and Goheung), and Jeolla-Do (Yeongam). 16S rRNA and MtCOI were amplified by PCR and sequenced for further phylogenetic analysis. In order to assess the route of spread of *B. tabaci* between countries, the 16S rRNA and MtCOI sequences of *B. tabaci* isolated from other countries (Japan, China, Sudan, Nigeria, Middle East Asia - Iran and Israel, and USA) were retrieved from Genbank (Supplementary Table 2). The sequence analysis data of 16S rRNA showed that *B. tabaci* was divided into two groups as shown for the TYLCV data. Group 1 contained the Korea isolates and those of Iran (AF247525), East Asia, Nigeria (AF247526), and Japan (AF236114 and AB293562) (Fig. 5A). Group 2 contained isolates from Sudan (AF272795), Egypt (AF246639), USA (AF246640), China (AF246642), the Dominican Republic (AF272798), and India (AF246643) (Fig. 5A).

In the case of the MtCOI sequences, sequences from eight China isolates, two Japan isolates, and one isolate each from Israel, Sudan, France, and Cameroon were obtained from Genbank (Supplementary Table 3). The phylogenetic analysis with *B. tabaci* MtCOI demonstrated that *B. tabaci* was also divided into two groups, similar to those for the 16S rRNA data (Fig. 5B). Group 1 contained the Korea isolates; other isolates from China Qinghai (GQ139503), Uruguay (EU760760), China Xinjiang (GQ139504), and France (EU760746) were closely aligned in the phylogenetic trees. Group 2 contained isolates from Pakistan Multan (FJ025788), Japan Okinawa (AB440792),

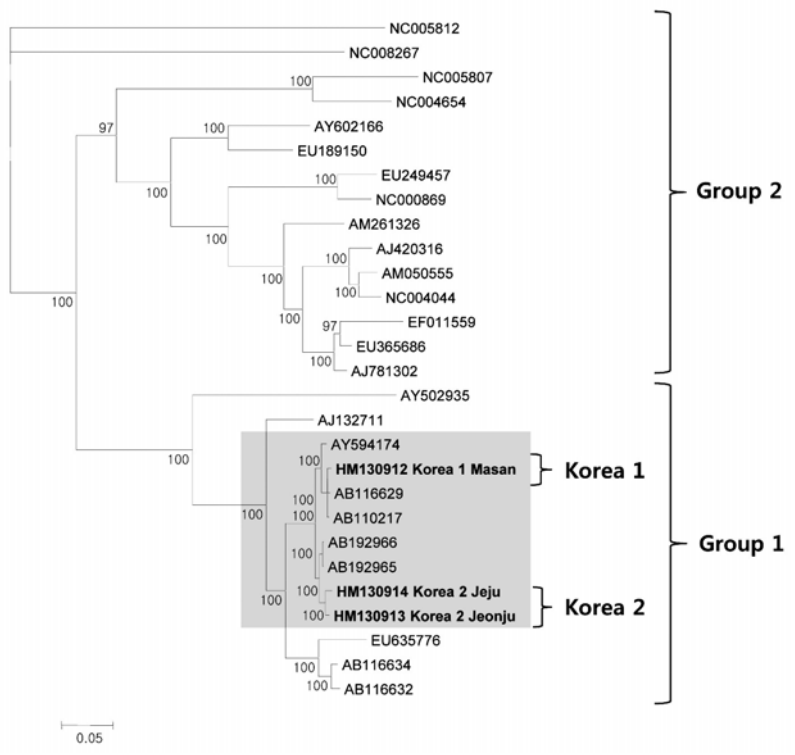

Fig. 4. Phylogenetic tree of TYLCV full sequence using Bayesian analysis. Full TYLCV sequence analysis employed the same sequences as in Fig. 3 coupled with the bayesian method. The total TYLCV sequence consisted of Groups 1 and 2; Group 1 included Korea 1 and 2 (Masan, Jeonju, and Jeju), and Group 2 included the other countries. The bayesian method was employed in the MCMC analysis of 10 million generations, with a cut off at 25% burn-in.

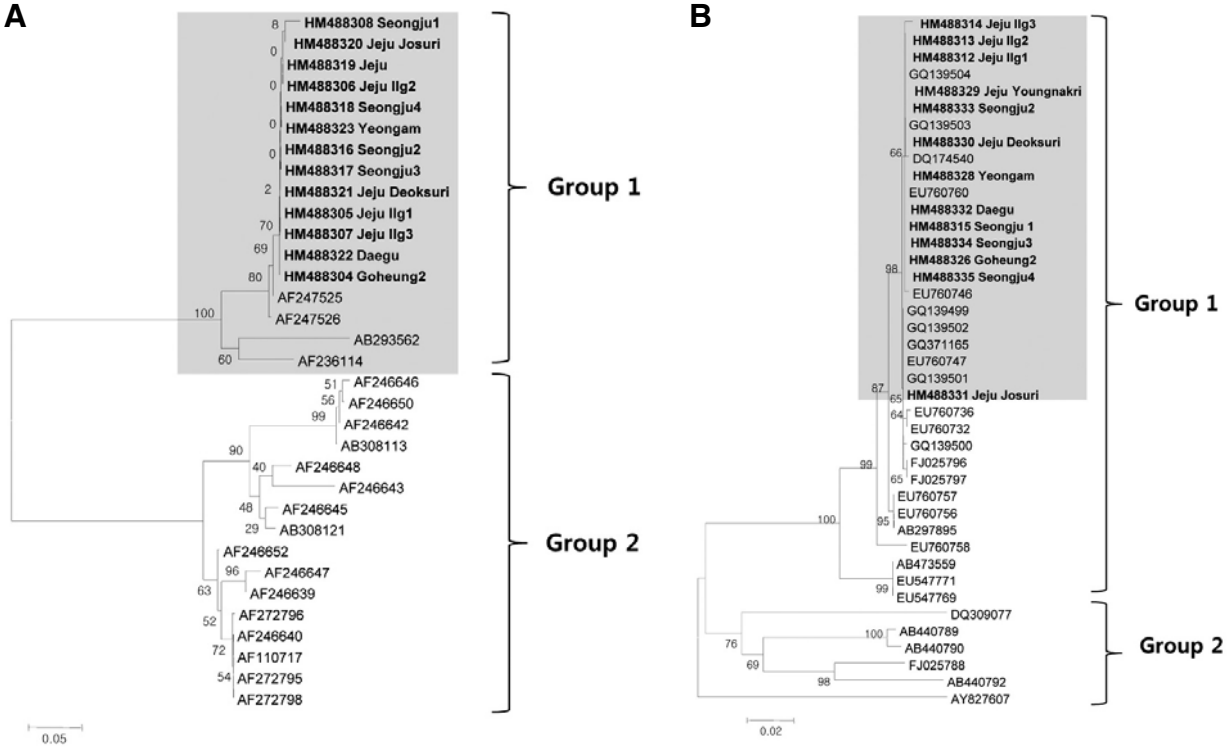

Fig. 5. Phylogenetic relationships of *B. tabaci* 16S rRNA and mitochondria cytochome oxidase (MtCOI) sequences. The sequences of *B. tabaci* 16S rRNA and MtCOI in isolates from various regions of Korea (Jeju, Yeongam, Daegu, and Seongju; 13 regions) were confirmed via the PCR analysis of genomic DNA and compared with sequences of isolates from other countries retrieved from Genbank. (A) Korea *B.tabaci* was closely aligned with the 16S rRNA sequences of the Iran (AF247525) and Nigeria (AF247526). (B) MtCOI sequences from Korea and others from China Qinghai (GQ139503), Uruguay (EU760760), China Xinjiang (GQ139504), and France (EU760746), were closely aligned in phylogenetic trees. Multi-alignments using the Muscle program and phylogenetic tree analysis using the MEGA 4.0 program in neighbor-joining algorithm, bootstrap of 1000 replicates, Scale bars represent 0.05 and 0.02.

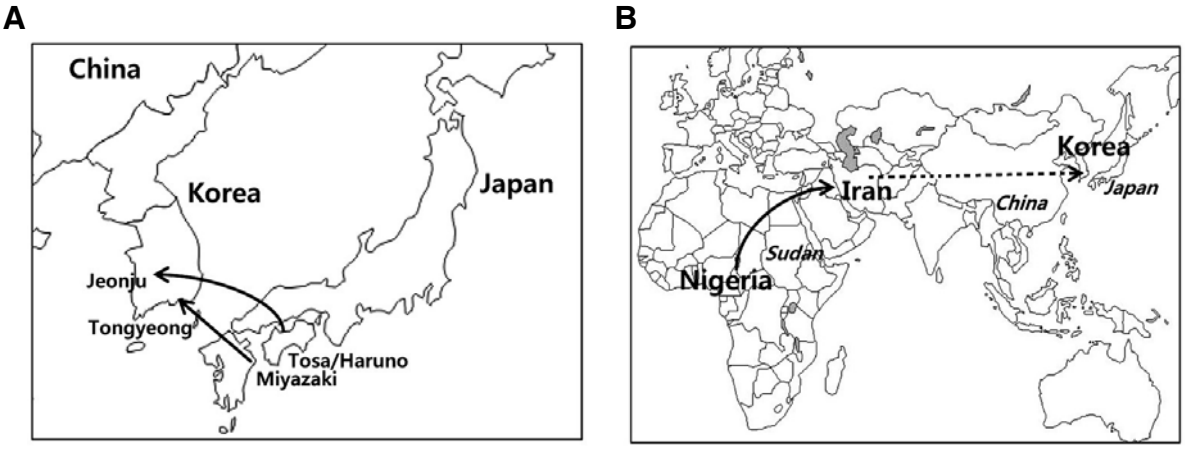

Fig. 6. Geographic map of East Asia and migration route of TYLCV and *B. tabaci*. (A) Phylogenetic tree results (ACI, AVI, and Full sequences) of TYLCV sequences demonstrated that the Korea sequences were closely related to TYLCV Japan from Miyazaki (AB116629) and Tosa (AB192965) /Haruno (AB192966) regions. (B) *B. tabaci* sequence data of 16S rRNA and MtCOI showed that Korea *B. tabaci* was closely aligned to *B. tabaci* of Iran (AF247525) and Nigeria (AF247526).

Japan Kagoshima (AB440790), and China (DQ309077) (Fig. 5B).

DISCUSSION

Full sequences of TYLCV were analyzed in tomatoes collected from Jeolla-Do (Jeonju), Gyeongsang-Do (Masan), and Jeju-Do (Jeju). TYLCV Korea isolates were regionally isolated and subgrouped into TYLCV Korea 1 group (2774-nucleotide) or TYLCV Korea 2 group (2781-nucleotide) (Fig. 2). When the nucleotide and amino acid sequences of each ORF and IR from the TYLCV Korea isolates and the TYLCVs from other countries were compared, all the amino acid sequences of the ORFs were found to be 96-100% similar to the Japan isolate (AB192965). Specifically, the amino acids of AV2 of the Korea and Japan isolates evidenced 100% identity (Table 2). Japan TYLCV originated from Israel, and therefore Korea TYLCV is also related to the Israel isolates with regard to sequence identity.

IR, a non-translational region located between AC1 and AV2, is known to be a highly variable region among *begomoviruses* (Yongping et al., 2008). Kumar et al. (2008) previously reported that the IR of *Tomato leaf curl Palampur virus* (ToLCPMV) evidenced an identity of only 54% with the ToLCV Taiwan (U88692), 84% with *Tomato leaf curl Cuba virus* (EU487042), 85% with the *Tomato leaf curl Philippines virus* (AB050597), and 55% with the *Tomato leaf curl Vietnam virus* (DQ641705). In some cases, the IR sequences varied even among identical hosts infected with the same virus (Kumar et al., 2008). TYLCV Korea 1 and 2 groups were also grouped according to IR sequence variations (Fig. 2B). A 7-base deletion was noted within a localized region of the IR of the group 1 isolates. The phylogenetic analysis and sequence comparison of AC1, AV1, IR, and the full sequences of TYLCVs isolated and sequenced in Korea with the TYLCV sequences of isolates from adjacent countries demonstrated that Korea TYLCV was generally and evolutionarily aligned with the Japan isolates (Figs. 3 and 4). However, the IR results showed that Korea TYLCV was closely aligned with ToLCV and HYVV (Fig. 3C). This can be explained by the genetic diversity in the IR sequences and recombination among the TYLCV Japan isolate, TbLCV (Shimizu and Ikegami, 1999) and HYVV (Ogawa et al., 2008) during the time at which TYLCV was introduced into Korea (Davino et al., 2009; Idris and Brown, 2005; Revill et al., 2003). Recombination events between viruses have been reported for a variety of DNA viruses, and increase the gene diversity of viruses through mutation, recombination, reassortment, and *de novo* acquisition (Padidam et al., 1999). There are two TYLCV groups in Japan: one is the Israel-originated TYLCV (AB192965) and the other is a TYLCV (AB192965) associated with relatively mild symptoms. On the basis of the sequence homology and phylogenetic analysis results, TYLCV Korea 1 group (Masan) and TYLCV Korea 2 group (Jeju/Jeonju) appear to originate from Israel, by way of Japan (Table 2, Figs. 3 and 4).

The TYLCV migration route was verified by the phylogenetic analysis of *B. tabaci*. The results of 16S rRNA analysis showed that Korea *B. tabaci* evolutionarily originated from Iran and Nigeria, again by way of Japan (Fig. 5A). Briddon et al. (2010) previously reported that geminiviruses evolved from the Old World (OW) to the New World (NW), and this hypothesis may explain the pattern of endemic TYLCV by a migration route of *B. tabaci* from Africa (Nigeria) to Japan via Middle Eastern Asia (Iran) (Fig. 5A). However, Korea *B. tabaci* was aligned at a further distance from Sudan *B. tabaci* than Nigeria *B. tabaci,* presumably because there was more trading traffic between Nigeria and Middle East Asia or Far East Asia. Therefore, the eggs or larvae of viruliferous *B. tabaci* were imported from Africa to Far East Asia via Middle East Asia (Padidam et al., 1995) (Figs. 6 and 7).

In this study, we demonstrate that TYLCV Korea was initially isolated in 2008 and rapidly became endemic. Based on the results of our phylogenetic analysis, TYLCV Korea can be classified into two groups designated as TYLCV Korea 1 and 2 groups, which originated in two different regions of Japan (Miyazaki and Tosa/Haruno) (Figs. 3 and 4). These TYLCV migration routes are supported by phylogenetic analysis of viruliferous *B. tabaci* (Figs. 4, 5, and 6). These data provide basic information in field of plant molecular virology regarding the genetic diversity of viruses as well as the interactions between plant viruses and their insect vectors.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

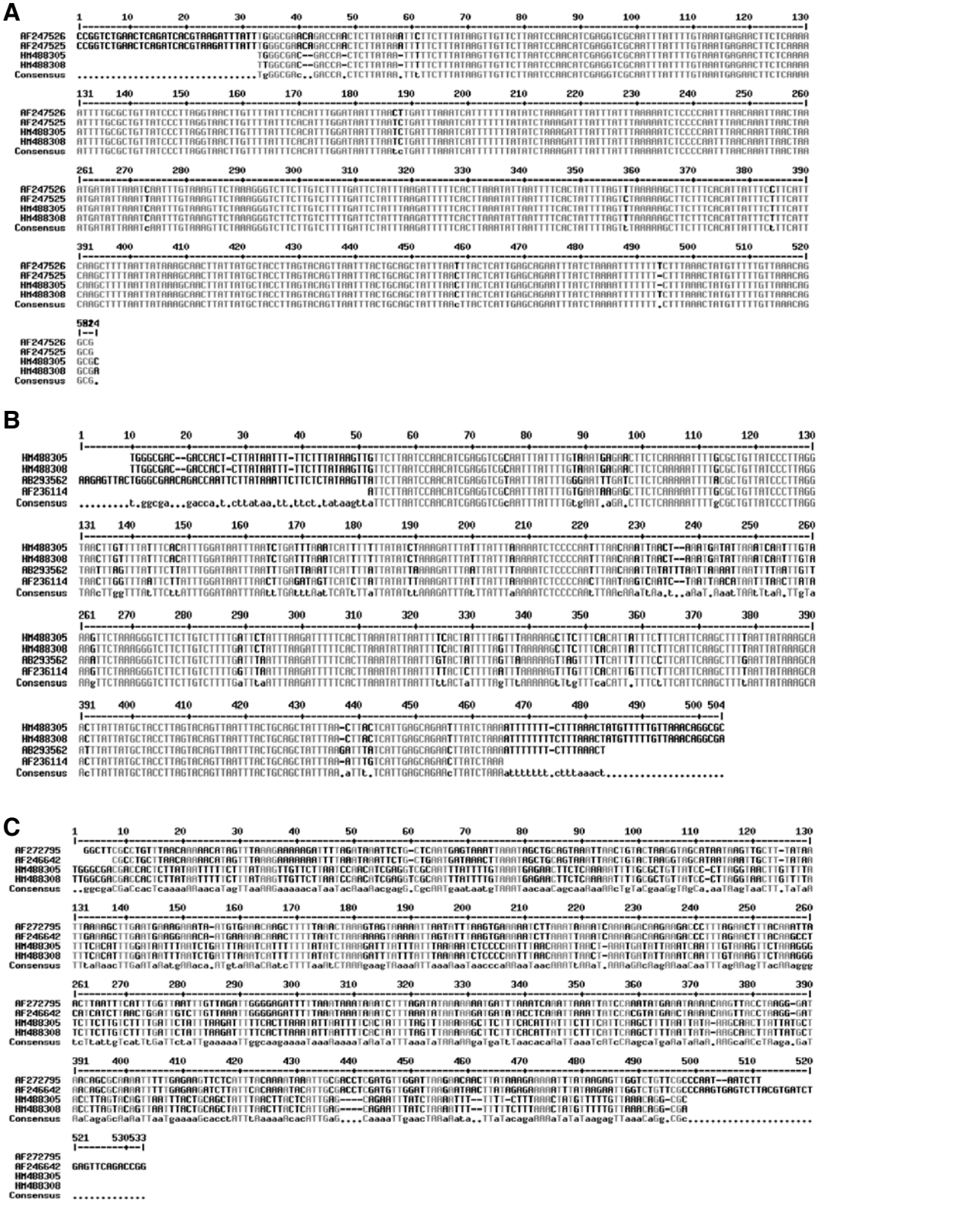

Fig. 7. Multi-alignment of *B. tabaci* 16S rRNA sequences. (A) *B. tabaci* was closely aligned to *B. tabaci* of Iran and Nigeria. Multi-alignment analysis showed that the *B. tabaci* 16S rRNA sequences of Iran (AF247525) and Nigeria (AF247526) were similar to the *B. tabaci* 16S rRNA sequences of Korea [Jeju Ilgwari (HM488305) and Seongju (HM488315)]. (B) Multi-alignment data of *B. tabaci* 16S rRNA sequences with two Japan lines (AB293562 and AF236114) and two Korea lines [Jeju Ilgwari (HM488305) and Seongju (HM488315)] showed more differences than those detected between Iran/Nigeria and Korea. (C) Sequence analysis based on *B. tabaci* 16S rRNA showed that the Korea *B. tabaci* may not have originated from the *B. tabaci* of Sudan (AF272795) and China (AF246642).

ACKNOWLEDGMENT

This research was supported by grant of the BioGreen21 Agenda program (No.PJ006325) from the Rural Development Administration in Korea.

REFERENCES

Boykin, L.M., Shatters, R.G., Jr., Rosell, R.C., McKenzie, C.L., Bagnall, R.A., De Barro, P., and Frohlich, D.R. (2007). Global relationships of *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. Mol. Phylogenet. Evol. *44*, 1306-1319.

Davino, S., Napoli, C., Dellacroce, C., Miozzi, L., Noris, E., Davino, M., and Accotto, G.P. (2009). Two new natural begomovirus recombinants associated with the tomato yellow leaf curl disease co-exist with parental viruses in tomato epidemics in Italy. Virus Res. *143*, 15-23.

De Barro, P.J., Driver, F., Trueman, J.W., and Curran, J. (2000). Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. Mol. Phylogenet. Evol. *16*, 29-36.

Dellaporta, S.L., Wood, J., and Hicks, J.B. (1983). A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. *1*, 19-21.

Everett, K.D., Thao, M., Horn, M., Dyszynski, G.E., and Baumann, P. (2005). Novel chlamydiae in whiteflies and scale insects: endosymbionts 'Candidatus Fritschea bemisiae' strain Falk and 'Candidatus Fritschea eriococci' strain Elm. Int. J. Syst. Evol. Microbiol. *55*, 1581-1587.

Fazeli, R., Heydarnejad, J., Massumi, H., Shaabanian, M., and Varsani, A. (2009). Genetic diversity and distribution of tomato-infecting begomoviruses in Iran. Virus Genes *38*, 311-319.

Ghanim, M., Morin, S., Zeidan, M., and Czosnek, H. (1998). Evidence for transovarial transmission of tomato yellow leaf curl virus by its vector, the whitefly *Bemisia tabaci*. Virology *240*, 295-303.

Hillis, D.M., and Dixon, M.T. (1991). Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol. *66*, 411-453.

Hsieh, C.H., Wang, C.H., and Ko, C.C. (2007). Evidence from molecular markers and population genetic analyses suggests recent invasions of the Western North Pacific region by biotypes B and Q of *Bemisia tabaci* (Gennadius). Environ. Entomol. *36*, 952-961.

Idris, A.M., and Brown, J.K. (2005). Evidence for interspecific-recombination for three monopartite begomoviral genomes associated with the tomato leaf curl disease from central Sudan. Arch. Virol. *150*, 1003-1012.

Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G.P., Crespi, S., and Gronenborn, B. (1991). Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. Nucleic Acids Res. *19*, 6763-6769.

Kim, H., and Lee, S. (2008). Molecular systematics of the genus Megoura (Hemiptera: Aphididae) using mitochondrial and nuclear DNA sequences. Mol. Cells *25*, 510-522.

Kim, M.I., Baek, J.Y., Kim, M.J., Jeong, H.C., Kim, K.G., Bae, C.H., Han, Y.S., Jin, B.R., and Kim, I. (2009). Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Mol. Cells *28*, 347-363.

Kumar, Y., Hallan, V., and Zaidi, A.A. (2008). Molecular characterization of a distinct bipartite begomovirus species infecting tomato in India. Virus Genes *37*, 425-431.

Lee, S., Stenger, D.C., Bisaro, D.M., and Davis, K.R. (1994). Identification of loci in Arabidopsis that confer resistance to geminivirus infection. Plant J. *6*, 525.

Luan, Y.S., Zhang, J., Liu, D.M., and Li, W.L. (2007). Molecular characterization of sweet potato leaf curl virus isolate from China (SPLCV-CN) and its phylogenetic relationship with other members of the Geminiviridae. Virus Genes *35*, 379-385.

Matsuda, N., Sharma, P., Bajet, N.B., and Ikegami, M. (2008). Molecular characterization of a new strain of tomato leaf curl Philippines virus and its associated satellite DNAbeta molecule: further evidence for natural recombination amongst begomoviruses. Arch. Virol. *153*, 961-967.

Navot, N., Pichersky, E., Zeidan, M., Zamir, D., and Czosnek, H. (1991). Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic component. Virology *185*, 151-161.

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

Padidam, M., Beachy, R.N., and Fauquet, C.M. (1995). Classification and identification of geminiviruses using sequence comparisons. J. Gen. Virol. *76*, 249-263.

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

Paredes-Esquivel, C., Donnelly, M.J., Harbach, R.E., and Townson, H. (2009). A molecular phylogeny of mosquitoes in the Anopheles barbirostris Subgroup reveals cryptic species: implications for identification of disease vectors. Mol. Phylogenet. Evol. *50*, 141-151.

Revill, P.A., Ha, C.V., Porchun, S.C., Vu, M.T., and Dale, J.L. (2003). The complete nucleotide sequence of two distinct geminiviruses infecting cucurbits in Vietnam. Arch. Virol. *148*, 1523-1541.

Sawangjit, S., Chatchawankanphanich, O., Chiemsombat, P., Attathom, T., Dale, J., and Attathom, S. (2005). Molecular characterization of tomato-infecting begomoviruses in Thailand. Virus Res. *109*, 1-8.

Shahid, M.S., Mansoor, S., and Briddon, R.W. (2007). Complete nucleotide sequences of cotton leaf curl Rajasthan virus and its associated DNA beta molecule infecting tomato. Arch. Virol. *152*, 2131-2134.

Shatters, R.G., Jr., Powell, C.A., Boykin, L.M., Liansheng, H., and McKenzie, C.L. (2009). Improved DNA barcoding method for *Bemisia tabaci* and related Aleyrodidae: development of universal and Bemisia tabaci biotype-specific mitochondrial cytochrome c oxidase I polymerase chain reaction primers. J. Econ. Entomol. *102*, 750-758.

Shimizu, S., and Ikegami, M. (1999). Complete nucleotide sequence and the genome organization of tobacco leaf curl geminivirus from Japan. Microbiol. Immunol. *43*, 989-992.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. *24*, 1596-1599.

Thao, M.L., Baumann, L., Hess, J.M., Falk, B.W., Ng, J.C., Gullan, P.J., and Baumann, P. (2003). Phylogenetic evidence for two new insect-associated Chlamydia of the family Simkaniaceae. Curr. Microbiol. *47*, 46-50.

Tsagkarakou, A., Tsigenopoulos, C.S., Gorman, K., Lagnel, J., and Bedford, I.D. (2007). Biotype status and genetic polymorphism of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Greece: mitochondrial DNA and microsatellites. Bull. Entomol. Res. *97*, 29-40.

Yongping, Z., Weimin, Z., Huimei, C., Yang, Q., Kun, S., Yanhui, W., Longying, Z., Li, Y., and Zhang, H. (2008). Molecular identification and the complete nucleotide sequence of TYLCV isolate from Shanghai of China. Virus Genes *36*, 547-551.

Zhang, H., Gong, H., and Zhou, X. (2009). Molecular characterization and pathogenicity of tomato yellow leaf curl virus in China. Virus Genes *39*, 249-255.