Genome characterization of *Zucchini yellow mosaic virus* infecting cucurbits reveals the presence of a new genotype in Trinidad & Tobago in the Caribbean region

Chinnadurai Chinnaraja
UWI: The University of the West Indies
Mounika Kollam
UWI: The University of the West Indies
Adesh Ramsubhag
UWI: The University of the West Indies
Jayaraman Jayaraj (jayauwi@gmail.com)
UWI: The University of the West Indies https://orcid.org/0000-0001-6961-2548

Original Article

Keywords: Zucchini yellow mosaic virus, ZYMV-Trini isolates, Virus recombination, ZYMV genotypes, Phylogenetic analysis, Aphis gossypii

Posted Date: February 10th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-203018/v1

Version of Record: A version of this preprint was published at Archives of Virology on April 3rd, 2021. See the published version at https://doi.org/10.1007/s00705-021-05048-4.

Abstract

Zucchini yellow mosaic virus is a *potyvirus*, which is becoming a serious pathogen of pumpkin and other cucurbits in Trinidad and Tobago and the entire Caribbean region. In this study, four *Zucchini yellow mosaic virus* (ZYMV) isolates infecting pumpkin in Trinidad and Tobago were characterized by complete genome sequencing for the first time. Phylogenetic analyses of the isolates showed variability of 5.9–6.0 % nt and 7.7–7.9 % aa sequences with the most closely related isolates NAT and AG (Israel) and SE04T (Slovakia). Based on the variations in complete genome as well as gene sequences, a new genotype designated ZYMV-Trini is proposed for these isolates. Among the gene sequences of ZYMV-Trini isolates, maximum variations were noticed in HC-Pro gene with 20.8 % aa sequence divergence from their closest relatives whereas the least variations were with the NIb, P3 and CP genes (1.8 to 2.2 % aa divergence). This study also proved that transmission of ZYMV can also occur through seeds (2 %), but this was less common than transmission via the aphid species *Aphis gossypii.* The progression of ZYMV in pumpkin seedlings was quantified by RT-qPCR which showed a rapid surge in viral load after 37 days. From the recombination analyses, it could be concluded that isolates SE04T from Slovakia, NAT from Israel and AG from Israel have major contributions in the genome architecture of ZYMV-Trini isolates.

1. Introduction

Zucchini yellow mosaic virus (ZYMV) is a member of the genus *Potyvirus* within the *Potyviridae* family. ZYMV was first reported in Italy in 1973 [27] and then subsequently spread worldwide, causing devastating epidemics in tropical, subtropical, and temperate regions. Members of the *Cucurbitaceae* are the primary hosts of ZYMV and disease symptoms include severe mosaic, yellowing, distortion of leaves, stunting in plant growth, severe fruit deformation and color cracking [8]. The symptoms can render fruits unmarketable and cause yield loss up to 94% [1, 19]. Transmission of ZYMV occurs both horizontally and vertically by aphids and seeds respectively, although horizontal transmission by aphids in a non-persistent manner is predominant [37]. Up to 26 species of aphids have been reported to transmit ZYMV experimentally, but only a few of them are commonly found in the field associated to transmission [18].

The ZYMV genome is ~ 9.6 kb in size and contains a positive sense single-stranded RNA molecule. The genome has one open reading frame (ORF) encoding for a single polyprotein precursor which is subsequently processed by three virally-encoded proteases into ten functional small mature proteins; P1 (protease), HC-Pro (helper component/protease), P3, 6K1, CI (cylindrical inclusion), 6K2, NIa (nuclear inclusion a), VPg (viral protein linked genome), NIb (nuclear inclusion b) and CP (coat protein) [26]. In addition, another short ORF has been found embedded within the P3 cistron (PIPO) which is translated in the + 2 reading frame [4]. The ZYMV's 5' untranslated region (UTR) is thought to contain two regulatory regions that are believed to direct cap independent translation [34] via interactions with the poly-A tail [11].

Cucurbits are major food crops of the Caribbean region, accounting for 27% of cultivated fields in Trinidad and Tobago with an average production of ~2,750 tons (pumpkin, squash, and gourds) per year (http://faostat3.fao.org/browse/Q/QC/E). The complete genome sequences of ZYMV isolates infecting cucurbits have been reported from several countries [7, 21, 26, 29, 41, 42] but not yet for the Caribbean region. A detailed survey conducted in pumpkin fields between 2014 and 2016 in six major cropping zones of Trinidad showed maximum ZYMV incidence (74 %) in dry season. Further, the infection of ZYMV along with the coinfection of *Squash mosaic virus* (SqMV) was also confirmed in cucurbits from Trinidad and Tobago [3].

In this study, the complete genome of ZYMV isolates from Trinidad and Tobago was sequenced for the first time. Phylogenetic and recombination analysis with the available ZYMV isolates from different geographical regions was carried out to study their evolutionary relationship and genetic diversity. The progression of ZYMV infection following aphid/seed transmission was also quantified.

2. Materials And Methods

2.1. Sample collection and RNA extraction

ZYMV infected pumpkin leaf samples collected from farmers' fields (10 from each location) from Barrackpore, Macoya, Las Lomas and Orange Groove in the island of Trinidad [3] were used for complete virus genome sequencing in this study. Total RNA was extracted from leaf samples (1 g) using TRI reagent (Sigma, USA) following the manufacturer's protocol.

2.2. RT-PCR and sequencing

Diagnosis of ZYMV was carried out by PCR using CP-forward (5'-GCTCCATACATAGCTGAGAC-3') and CP-reverse

(5'-AACGGAGTCTAATCTCGAGC-3') primers targeting a partial coat protein region (1100-bp) of ZYMV. Ten different sets of primers targeting overlapping fragments of ZYMV-polyprotein were designed (Fig. 1; Supplementary Table 1) to derive the complete genome of the Trinidad isolates. ImProm-II™ Reverse Transcription System (Promega, USA) was used for the synthesis of complementary DNA using 1 µg of RNA. PCR reactions were performed in a thermocycler (Techne, USA) for all primer pairs. Each PCR reaction (25µL) contained 100 ng of cDNA, 1 unit of *Pfu* DNA Polymerase, 10X Buffer with $MgSO_4$, 0.5 µl of 10 mM dNTP mix, 50 pmol of primer pairs and sterile milliQ water to the final volume. The PCR conditions were, an initial denaturation at 94 °C for 4 min; 30 cycles of 94 °C for 1 min, 54–60 °C for 1 min, 72 °C for 1 min; and a final primer extension step for 10 min at 72 °C. Amplicons were visualized by electrophoresis on 1.5 % agarose gels stained with ethidium bromide. Three replicates of all the PCR amplicons were gel purified using Gen-Elute Gel Extraction Kit (Sigma, USA) and cloned in pGEM®T vector (Promega, USA) before Sanger sequencing (Macrogen Inc, USA) of both strands. Complete genomes of ZYMV isolates from Trinidad were constructed by aligning all partial overlapping fragments of ZYMV polyprotein sequences with reference genome sequences of ZYMV from NCBI-GenBank using bioinformatics software Mega X [20].

2.3. Phylogenetic analysis

All the nucleotide (nt) and amino acid (aa) sequence fragments targeting the polyprotein of ZYMV were separately aligned with reference sequences from GenBank (Supplementary Table 2) using Clustal W and MAFFT [22, 29]. The Genome Annotation Transfer Utility (GATU) [38] was then used to annotate the complete genome sequences of the Trinidad isolates using the reference isolate TW-TN3 (AF127929.2) obtained from RefSeq database. Phylogenetic analysis was carried out with the ZYMV Trinidad isolates and 63 complete genome sequences obtained from GenBank representing different geographical regions of the world (Supplementary Table 2). Maximum Likelihood trees were generated using the Tamura-Nei model for complete nt genome sequences of ZYMV and aa sequences of individual genes P1, HC-Pro, P3, CI, NIb and CP in Mega X software [20] with 1000 bootstrap replications. Similarity matrices revealing the percentage identities among nt and aa sequences of all the clusters in phylograms were also generated in Mega X.

2.4. Aphid transmission and RT-qPCR quantification of the virus

In order to confirm the transmission of ZYMV through aphid vectors, sterile pumpkin seedlings were grown in a greenhouse. Single adult aphids (*Aphis gossypii*) from a virus-free colony were transferred to pumpkin seedlings infected with ZYMV (and confirmed by PCR reactions) for acquisition feeding for 48 h. The viruliferous aphids were then transferred on to 15 sterile pumpkin seedlings individually for 48 h for inoculation feeding in a netted greenhouse box. After inoculation, the aphids were killed using malathion treatment. After seven days, the total RNA was extracted from leaf samples (1 g of 3rd leaf) of inoculated and control seedlings using TRI reagent (Sigma, USA). The RNA was reverse transcribed and PCR amplification of the cDNA carried out using ZYMV-specific primers (CR-for/CP-rev) as before. Amplification of a 1,100-bp fragment in 12 out of 15 receptor seedlings confirmed the presence and transmission of ZYMV in pumpkin. The PCR products were subjected to Sanger sequencing for cross checking identity by BLAST (NCBI) which validated infection by ZYMV. Leaf samples from 10 out of 12 ZYMV positive seedlings were collected every 10 days, after

the initial sampling, until flowering and RNA was extracted. The virus titres were quantified from 10 samples with three replicates by RT-qPCR to assess the progression of infection at different growth stages.

In addition, 40 fruit samples were collected from 40 different ZYMV PCR-positive pumpkin plants from the field. All seeds were separated from the fruits and surface sterilized in 70% ethanol for one minute and 5% sodium hypochlorite for 5 minutes and then washed 4 times with distilled water to ensure safe removal of all contaminants. All the surface sterilized seeds were pooled together and one-hundred seeds were randomly collected and planted in individual pots. Leaf samples were collected after 7 days of germination. PCR reactions confirmed ZYMV infection in 2 out of 100 seedlings raised from the seeds. Leaf samples from those 2 positive seedlings were collected every 10 days thereafter for viral quantification as before.

For cDNA synthesis, RNA (500 ng) was reverse transcribed using the Multiscribe Reverse Transcriptase Kit (Invitrogen, USA), primed with 40 nmol of ZYMVRT- R1 (5'- GGCCAAACAACCTTGAAGAAACATTGC – 3') primer in a 20-µL reaction following the manufacturer protocol using thermocycler (Techne, USA). Real-time quantitative PCR was performed with three replicates of each sample with 500 ng of cDNA template, 12.5 uL of SYBR^{R} Green JumpStart™ *Taq* ReadyMix™ (Sigma, USA), 50 nM of primer pairs ZYMVRT- F1 (5'- GAGAAATGCAGAGGCACCATACATGCCG – 3') and ZYMVRT- R1 (5'- GGCCAAACAACCTTGAAGAAACATTGC – 3') targeting a 181 nt region of the ZYMV coat protein gene. RT-qPCR (25 uL) was carried out in an Applied Biosystems 7500 Fast Real Time PCR system (Life Technologies Corp., USA). A region of the 18S rRNA gene was used as an endogenous control in all the samples. The optimised RT-qPCR conditions were; 95 °C for 10 min; 40 cycles of 95 °C for 15 sec and 60 °C for 1 min. Melt curve analysis was performed at 60 °C for 15 sec to ensure that a homogenous amplification product was produced. ZYMV infected samples from greenhouse pots were used as a positive control.

Relative standard curve analysis was done using 500 ng of cDNA from the ZYMV-positive control; a 10-fold serial dilution was performed to generate the standard curve. The threshold cycle number (CT) of each dilution were plotted against standard concentrations of cDNA and standard curve was constructed. The regression line generated in the standard curve was used to quantify the titre of ZYMV from the respective samples at each sampling time.

In addition, expression level of ZYMV-CP target was assessed by relative quantification method using $2^{-(\Delta\Delta Ct)}$ values. ΔCt was derived by subtracting Ct of endogenous control (18S rRNA) from Ct of test samples (ZYMV-CP). ΔΔCt was derived by subtracting ΔCt of negative control from ΔCT of samples. Greenhouse grown pumpkin plants were used as positive and negative control (PCR confirmed) for the qPCR analysis.

2.5. Recombination analysis

Recombination analysis was performed with all the 63 complete genome nt sequences of ZYMV isolates (Supplementary Table 2) from different geographical regions around the world to detect the presence of recombination sites using the RDP, GENECONV, BOOTSCAN, MAXIMUM CHISQUARE, CHIMAERA, SISTER SCAN and 3SEQ non-parametric recombination detection methods as implemented in RDP5 software [31, 32]. A multiple comparison corrected P-value cut-off of 0.05 and default settings were used throughout the analysis, and only events detectable with three or more different methods were retained for further analysis.

3. Results

RT-PCR analysis with the diagnostic primers CP-for/ CP-rev confirmed ZYMV infection in all the samples collected from Barrackpore, Macoya, Las Lomas and Orange Groove. All fragments of ZYMV-polyprotein were amplified specifically using newly designed primer pairs (Supplementary Table 1) and sequenced. Overlapping sequences were aligned with multiple reference isolates collected from the GenBank (Supplementary Table 2) and the complete genome was constructed for four Trinidad isolates that were designated as ZYMV-Trini isolates. The complete genome of the four ZYMV-Trini isolates

was 9594 nts long, encoding a polyprotein of 3,081 aa residues comprising 10 different genes. The genes encoded in polyprotein of ZYMV-Trini isolates were P1 (930 nt), HC-Pro (1368 nt), P3 (1038 nt), 6K1 (156 nt), CI (1902 nt), 6K2 (159 nt), VPg (570 nt), NIa (729 nt), NIb (1551 nt), Coat protein (CP) (837 nt) and PIPO (201 nt). The 5' and 3' UTR regions were 139 and 212 nt in size, respectively. The complete nucleotide sequences of all four new Trinidad isolates were deposited in GenBank under accession numbers as ZYMV-Trini1 (MF072712), ZYMV-Trini2 (MF072713), ZYMV-Trini3 (MF072714) and ZYMV-Trini4 (MF072715).

The pairwise identities among the nt sequences of ZYMV-Trini isolates were 99.9-100 %. A phylogram constructed based on the complete genome nt sequences with 63 reference isolates from around the world showed that ZYMV-Trini isolates formed a separate cluster (Figure 2; Table 1). The ZYMV-Trini isolates were most closely related (nt similarity = 94.0 to 94.1) to reference isolates that included NAT (Israel), AG (Israel) and SE04T (Slovakia). The polyprotein region of ZYMV-Trini isolates showed 93 variable amino acids with their closest relatives; NAT (Israel), AG (Israel) and SE04T (Slovakia) (Supplementary Table 3, 4). The isolates Per-1 (Australia), Knx-25 (Australia) and ZYMP13PREP (Reunion Island) were found to have the least aa similarities (23.2-23.6 % nt variations) with the ZYMV-Trini isolates (Table 1).

Phylogenetic analyses of the aa gene sequences of the ZYMV-Trini and reference isolates showed the four local isolates formed a separate and distinct cluster for the HC-Pro, CI, and NIb genes (Supplementary Figure 1). Pairwise comparison with all the reference P1 genes showed the Trini sequences were closely related to isolate TW-TN3 (Taiwan) with 97.6 to 97.7 % nt identities. Similar analyses for the HC-Pro and NIb genes showed that Z5-1 (Japan) was the most closely related isolate to the ZYMV-Trini isolates with nt similarities of 90.7 % and 94.6 %, respectively. Pairwise comparison with reference sequences for the P3, CI and CP genes showed that the Trini isolates were highly related to the isolate Z-104 (Italy) with nt similarities of 97.6, 98.0 and 97.2 %, respectively (Table 1; Figure 2, Supplementary Figure 1).

The isolate ZYMP13PREP (Reunion Island) was found to have the lowest nt similarity to the Trini isolates for the genes CI (81.4 %) and NIb (82.9 %). Whereas, the isolate WM (China) showed the least nt similarity with CP gene (82.7 %) and the isolate Singapore (Singapore) showed the least similarity with the P1 (62.8-62.9 %) and HC-Pro (80.7 %) gene sequences (Table 1).

3.1. Progression of virus titre

The transmission of ZYMV through aphids was confirmed through PCR in 12 out of 15 inoculated pumpkin seedlings. Subsequently, those 10 positive seedlings were used for the quantification of ZYMV-CP targets by qPCR. In the case of seed transmission, only 2% of seedlings (2 out of 100 seedlings) were found positive by diagnostic PCR reactions which were further analysed for qPCR quantification.

A standard curve generated using known quantities of cDNA from the positive control and plotted against the threshold cycles (CT) values of each dilution resulted in the linear equation $y = 4.034x + 12.066$ and correlation, $r^2 = 0.9927$. cDNA quantities representing the progression of ZYMV-RNA targets for the Ct mean of all samples with three replicates (10 different seedlings in aphid transmission studies and two different seedlings in seed transmission studies) collected in 10 days' intervals were quantified using the standard curve and, the mean values are represented in Figure 3. Melt curve analysis also confirmed the specificity of the primers in RT-qPCR.

Relative gene expression studies based on $2^{-(\Delta\Delta Ct)}$ showed that the titre of ZYMV-CP were minimum in samples collected between 7 to 27 days and it increased rapidly after 37 days. The maximum titre was observed after 50 days in all the plant samples after aphid and transmission (Table 3).

3.2. Recombination analysis

Phylogenetic analysis of ZYMV based on the complete genome nt sequences revealed differences to the maximum of 23.6 % among isolates reported around the world. In silico recombination analyses using seven different detection methods with all the isolates in this study detected recombination sites throughout the genome of ZYMV. Twelve recombination sites were detected in eight different isolates from seven countries (Figure 4).

In ZYMV-Trini isolates, the first recombination site was detected with partial P3, complete 6K1, C1, 6K2, VPg and partial NIa gene sequences, with the isolates SE04T (Slovakia), NAT (Israel), AG (Israel) as a major parents and WM (China) as a minor parent. The second recombination site in ZYMV-Trini isolates was found with partial P1 gene sequences and isolate SB-02 (India) found to be a major parent. The hypothetical parental and daughter sequences which would fix the patterns of recombination sites were assigned using seven non-parametric methods as implemented in RDP5 software. Isolates ZYMV-WS from China, ZYMV-Trini from Trinidad, and Z-104 from Italy recorded two recombination sites in their genome. It was noteworthy that ZYMV-Trini isolates were found as a major parent for three different isolates Z5-1 (Japan), Z-104 (Italy) and ZYMV-WS (China), and also as a minor parent for ZYMV-WS (China) for its second recombination site (Table 2).

4. Discussion

ZYMV is becoming a serious pathogen in most cucurbit growing regions of the world where the infection rates of at least 40% has been reported in tropical and sub-tropical regions. Disease surveys in pumpkin have indicated ZYMV disease incidence levels reported up to 75 % in dry season in Trinidad [3].

Natural populations of RNA viruses rapidly generate genetic diversity because of a combination of high mutation rates, rapid replication, recombination events, high frequency of occurrence and a variety of strains [9]. In this study, phylogenetic analysis revealed that Trinidad isolates form a new genotype ‘ZYMV-Trini’ since they have variations of 5.9-6.0 % in complete nt sequences as compared to their closest known relatives, including isolates NAT and AG (Israel) and SE04T (Slovakia). The within-virus species genotype classification system adopted a neutral nomenclature involving letters of the alphabet and Latinized numerals that avoid potentially misleading names [29, 30]. Phylograms of aa sequences also showed the ZYMV-Trini isolates forming a separate cluster for HC-Pro, CI, and NIb genes in this study. In case of the phylogram of the P1 gene sequences, the isolate TW-TN3 (Taiwan) was closely related to the Trini isolates, with 2.3–2.4 % nt variation level.

The capsid protein (CP) gene of potyviruses is widely used as a valid typing tool to differentiate among isolates [36]. However, comparison of complete genome sequences allows for a more complex analysis of virus variability and may provide information on the evolutionary history and existence of major evolutionary events, such as recombination, as was demonstrated for various potyviruses [14, 41]. Among the gene sequences in the polyprotein of ZYMV-Trini isolates, NIb and CP are highly conserved but the HC-Pro gene had the maximum variation in aa sequences, as compared to the closely related isolates. In a phylogram constructed based on coat protein aa sequences, all previously reported ZYMV-coat protein sequences from Trinidad and Tobago and ZYMV-Trini isolates from this study get grouped together and this may suggest that all the ZYMV isolates from Trinidad and Tobago belongs to the same genotype (data not given). Among the other reference isolates, ZYMP13PREP from Reunion Island had maximum variations from the Trini isolates, viz., 23.6 % with complete nt genome, and 18.6 % with CI and 7.1 % with NIb aa sequences (Table 1).

Aphids are the most successful vectors of potyviruses, due to an array of generic and specific features they possess [15], including precision delivery of viral particles via the stylet, parthenogenetic mode of reproduction within a short span of time, diverse range of host plants, survival in adverse conditions and the ability to disseminate over long distances [28, 33]. Katis et al. [18] reported the most abundant aphid vectors of ZYMV in a study in Greece included *M. persicae, Aphis gossypii* and *Aphis spiraecola*. We also reported *A. gossypii* as a vector of ZYMV in Trinidad in our earlier preliminary study [3] and it was reconfirmed in the current study. The relative standard curve method through RT-qPCR effectively

detected the incremental increase of ZYMV-RNA targets in pumpkin seedlings following transmission through *A. gossypii*. Vector transmission occurs as a result of interaction between the aphid stylet, and viral proteins of ZYMV such as coat protein (CP) and helper component proteinase (HC-Pro). Specifically, the DAG motif on the CP interacts with the PTK region of the HC-Pro, and a secondary motif on the HC-Pro (KLSC) interacts with the aphid stylet [40]. Volunteer cucurbitaceous crop and weed plants also act as infection sources for ZYMV spread to cucurbit crops [5, 6, 23].

Generally, ~ 20% of viral plant pathogens are known to be seed-transmitted. Seed to seedling transmission rate was earlier reported for ZYMV at a low level (1.6%) [17, 37]. A similar trend was also observed in this study, with only 2% seed transmission rate. The vertical transmission of ZYMV by seed is less common than horizontal transmission via aphids and also many studies support the hypothesis that the insect vector is an important factor in inducing virus variation [35]. In both aphid and seed transmission experiments, ZYMV increments in the plant were steady up to 37 days but a rapid surge was observed by RT-qPCR analysis between 37 and 57 days in this study.

Through recombination, viruses gain pathogenicity or virulence, and the ability to invade new hosts [13, 16]. Recombination is advantageous for RNA viruses as it can create high fitness genotypes more rapidly than mutation alone [2]. This study also supports the hypothesis that recombination is a dominant feature of ZYMV evolution as in other RNA viruses. Recombination sites detected *in silico* using RDP5 software suggested that the entire ZYMV genome is prone to recombination, though hotspots are concentrated in P1, HC-Pro, P3 and CP gene sequences in several isolates. Recombination breakpoints in the ZYMV-Trini isolates were noticed in P1 and between P3 and Nlb gene sequences. Maina et al. [29] earlier reported the same pattern of recombination breakpoints in ZYMV populations from East Timor and northern Australian cucurbit crops. Natural recombinants may emerge in virus populations only if they maintain relatively good fitness that includes preserving the functionality of each viral protein and the functional interactions between proteins [25, 31]. For plant viruses, the recombination rate might be much higher than expected, whereas rates for potyviruses may be up to ~ 25 % although only a small fraction of the generated variants emerge in the population due to strong selection pressure [10].

Isolates SE04T (Slovakia), AG (Israel) and NAT (Israel) were determined to be major parental contributors, circumstantially, for the genome architecture of ZYMV-Trini isolates. These parental isolates have 98.0-98.4 % nt identities among themselves and 93.4–94.1 % similarities with ZYMV-Trini isolates. Cucurbit cultivation in the Caribbean islands including Trinidad and Tobago is mainly dependent on imported seeds from different countries, and seed producing countries such as Israel, and China play a significant role in seed transfer. ZYMV can also move to new locations in ZYMV-infected fruit from which aphids can acquire and spread the virus [24]. Introductions could also occur from migrating birds carrying virus-infected seed in their intestines or discarded infected cucurbit fruit left behind by fishermen from neighbouring countries camping on the shore [12]. Maina et al. [29] reported the absence of genetic connectivity between ZYMV sequences from Papua New Guinea (PNG) and those from Australian or East Timor. The highest nucleotide similarity between a ZYMV sequence from PNG and elsewhere was 92.8% and the authors suggested the divergence could be due to a single introduction of ZYMV into PNG with subsequent evolution to adapt in this new environment.

It is also interesting to know that ZYMV-Trini isolates may have contributed as a major parent for isolates such as Z5-1 (Japan), Z-104 (Italy) and ZYMV-WS (China) for various genome fragments. However, more data need to be generated from the Caribbean island countries to study the genome dynamics of ZYMV-Trini isolates and their genetic connectivity among the isolates from neighbouring countries.

This study provides the first report of the complete nucleotide sequence of ZYMV from Trinidad and Tobago and further also highlights that recombination is a major driving force in the evolution and emergence of new variants of ZYMV. It is also noteworthy to understand the complexity of the variability of ZYMV isolates in order to derive effective field control measures.

Declarations

Acknowledgments

The authors are grateful to ACP-EU (Grant No. FED/2011/281 – 139) for the funding and support to carry out the research work. Logistical assistance provided by the collaborating institutions and Agriculture Ministries of the Caribbean states is duly acknowledged. Thanks for the ACP-EU project research group from the Department of Life Sciences, UWI, St. Augustine, Trinidad and Tobago for assistance in plant sampling.

References

1. Blua MJ, Perring TM (1989) Effect of zucchini yellow mosaic-virus on development and yield of cantaloupe (Cucumis-Melo). Plant Dis 734:317–320.
2. Chare ER, Gould EA, Holmes EC (2003) Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. J Gen Virol 84:2691–2703.
3. Chinnaraja C, Ramkissoon A, Rajendran R, Tony ST, RamsubhagA, Jayaraj J (2016) First Report of *Zucchini yellow mosaic virus* and *Squash mosaic virus* Infecting Cucurbits in Trinidad. Plant Dis 100(4):866.
4. Chung BY, Miller WA, Atkins JF, Firth AE (2008) An overlapping essential gene in the Potyviridae. Proc Natl Acad Sci 105:5897-5902.
5. Coutts BA, Jones RAC (2005) Incidence and distribution of viruses infecting cucurbit crops in the Northern Territory and Western Australia. Crop Pasture Sci 56:847-858.
6. Coutts BA, Kehoe MA, Webster CG, Wylie SJ, Jones RAC (2011) Zucchini yellow mosaic virus: Biological properties, detection procedures and comparison of coat protein gene sequences. Arch Virol 156:2119-2131.
7. Desbiez C, Gal-On A, Girard M, Wipf-Scheibel C, Lecoq H (2003) Increase in Zucchini yellow mosaic virus symptom severity in tolerant zucchini cultivars is related to a point mutation in P3 protein and is associated with a loss of relative fitness on susceptible plants. Phytopathology 93:1478-1484.
8. Desbiez C, Lecoq H (1997) Zucchini yellow mosaic virus. Plant Pathol 46:809–829.
9. Duffy S, Shackelton LA, Holmes EC (2008) Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 94:267-276.
10. Froissart R, Roze D, Uzest M, Galibert L, Blanc S, Michalakis Y (2005) Recombination very day: abundant recombination in a virus during a single multi-cellular host infection. PLoS Biol 3:389–395.
11. Gallie DR (2001) Cap-independent translation conferred by the 5’ leader of tobacco etch virus is eukaryotic initiation factor 4G dependent. J Virol 7524:12141–12152.
12. Gibbs AJ, Mackenzie AM, Wei K, Gibbs MJ (2008) The potyviruses of Australia. Arch Virol 153:1411-1420.
13. Gibbs AJ, Ohshima K (2010) Potyviruses and the digital revolution. Ann Rev Phytopathol 48:205-223.
14. Glasa M, Palkovics L, Kominek P, Labonne G, Pittnerova S, Kudela O, Candresse T, Subr Z (2004) Geographically and temporally distant natural recombinant isolates of *Plum pox virus* (PPV) are genetically very similar and form a unique PPV subgroup. J Gen Virol 85:2671–2681.
15. Harris KF, Maramorosch K (1997) Aphids as Virus Vectors, first ed. Academic Press, Cambridge. ISBN 978-14-8327-388-4.
16. James D, Sanderson D, Varga A, Sheveleva A, Chirkov S (2016) Genome sequence analysis of new isolates of the Winona strain of *Plum pox virus* and the first definitive evidence of intrastrain recombination events. Phytopathology 106:407-416.

17. Johansen E, Edwards MC, Hampton RO (1994) Seed transmission of viruses: current perspectives. Annual review of phytopathology 32:363–386.
18. Katis NI, Tsitsipis JA, Lykouressis DP, Papapanayotou A, Margaritopoulos JT, Kokinis GM, Perdikis DC, Manoussopoulos IN (2006) Transmission of zucchini yellow mosaic virus by colonizing and noncolonizing aphids in Greece and new aphid species vectors of the virus. J Phytopath 1545:293–302.
19. Khanal V, Ali A (2019) Complete genome sequence of a Zucchini yellow mosaic virus isolated from pumpkin in Oklahoma. Microbiol Resour Announc 82:e01583-18.
20. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547-1549.
21. Kwon SW, Kim MS, Choi HS, Kim KH (2005) Biological characteristics and nucleotide sequences of three Korean isolates of *Zucchini yellow mosaic virus*. J Gen Plant Pathol 71:80-85.
22. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 20. Bioinformatics 23:2947-2948.
23. Lecoq H, Desbiez C (2008) *Watermelon mosaic virus* and *Zucchini yellow mosaic virus* Encyclopedia Virol 5:433-440.
24. Lecoq H, Desbiez C, Wipf-Scheibel C, Girard M (2003) Potential involvement of melon fruit in the long-distance dissemination of cucurbit potyviruses. Plant Dis 87:955-959.
25. Lefeuvre P, Lett JM, Reynaud B, Martin DP (2007) Avoidance of protein fold disruption in natural virus recombinants. PLoS Pathogens 3:1782–1789.
26. Lin SS, Hou RF, Yeh SD (2001) Complete genome sequence and genetic organization of a Taiwan isolate of *Zucchini yellow mosaic virus*. Bot Bull Acad Sin 424:243-250.
27. Lisa V, Boccardo G, D'Agostino G, Dellavalle G, D'Aquilio M (1981) Characterization of a potyvirus that causes zucchini yellow mosaic. Phytopathology 71:667-672.
28. Lovisolo O, Hull R, Rosler O (2003) Coevolution of viruses with hosts and vectors and possible palaeontology. Adv Virus Res 62:325–379.
29. Maina S, Barbetti MJ, Edwards OR, Minemba D, Areke MW, Jones RAC (2019) *Zucchini yellow mosaic virus* Genomic Sequences from Papua New Guinea: Lack of Genetic Connectivity with Northern Australian or East Timorese Genomes, and New Recombination Findings. Plant Dis 103:1326-1336.
30. Jones RA, Kehoe MA (2016) A proposal to rationalize within-species plant virus nomenclature: Benefits and implications of inaction. Arch Virol 161(7):2051-2057.
31. Martin DP, Varsani A, Roumagnac P, Botha G, Maslamoney S, Schwab T, Kelz Z, Kumar V, Murrell B (2020) RDP5: A computer program for analysing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evolution, veaa087, https://doi.org/10.1093/ve/veaa087.
32. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis of recombination patterns in virus genomes, Virus Evolution 1(1), vev003, https://doi.org/10.1093/ve/vev003.
33. Ng JC, Perry KL (2004) Transmission of plant viruses by aphid vectors. Mol Plant Pathol 5:505–511.
34. Niepel M, Gallie DR (1999) Identification and characterization of the functional elements within the tobacco etch virus 5’ leader required for cap-independent translation. J Virol 7311:9080–9088.
35. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3:336–340.
36. Rybicki EP, Shukla DD (1992) Coat protein phylogeny and systematics of potyviruses. Arch Virol Suppl 5:139-170.
37. Simmons HE, Holmes EC, Gildow FE, Bothe-Goralczyk MA, Stephenson AG (2011) Experimental verification of seed transmission of Zucchini yellow mosaic virus. Plant Dis 956:751–754.

38. Tcherepanov V, Ehlers A, Upton C (2006) Genome Annotation Transfer Utility (GATU): rapid annotation of viral genomes using a closely related reference genome. BMC Genomics 7:150. doi:10.1186/1471-2164-7-150
39. Tomimura K, Gibbs AJ, Jenner CE, Walsh JA, Ohshima K (2003) The phylogeny of Turnip mosaic virus: comparisons of 38 genomic sequences reveal a Eurasian origin and a recent “emergence” in east Asia. Mol Ecol 128:2099-2111.
40. Urcuqui-Inchima S, Haenni AL, Bernardi F (2001) Potyvirus proteins: a wealth of functions. Virus Research 74(1-2):157–175.
41. Wisler GC, Purcifull DE, Hiebert E (1995) Characterization of the P1 protein and coding region of the *Zucchini yellow mosaic virus*. J Gen Virol. 76:37-45.
42. Zhao MF, Chen J, Zheng HY, Adams MJ, Chen JP (2003) Molecular analysis of *Zucchini yellow mosaic virus* Isolates from Hangzhou, China. Journal of Phytopathology 151:307-311.

Tables

Table 1 Nucleotide (Nt) and Amino acid (Aa) identities (%) of ZYMV- Trini isolates with other clusters in phylograms

Gene/ Genome	Cluster 1 (Nt/Aa)	Cluster 2 (Nt/Aa)	Cluster 3 (Nt/Aa)	Cluster 4 (Nt/Aa)	Cluster 5 (Nt/Aa)	Cluster 6 (Nt/Aa)	Cluster 7 (Nt/Aa)	Cluster 8 (Nt/Aa)	Cluster 9 (Nt/Aa)
Complete genome	99.9-100 / 99.9-100	93.1-94.1/ 90.7-92.3	93.6-93.7/ 91.9-92.0	93.6-93.7/ 91.6-91.7	91.0/86.6	92.1-92.8/ 91.2-91.7	90.8-91.9/ 88.7-89.4	89.4-90.8/ 88.3-90.0	76.4-78.2/ 77.3-79.0
P1	85.0-89.4 / 71.0-77.6	86.6-88.2 / 72.0-76.4	95.6-100 / 91.6-100	92.1-92.4 / 83.9-86.6	90.7-92.3 / 81.7-85.3	88.9-89.7 / 77.2-79.5	62.9-65.3 / 52.4-54.6	-	-
HC-Pro	89.0-90.2 / 73.1-75.3	85.9-90.3 / 70.4-75.7	90.7-90.8 / 74.7-74.9	89.9-90.5 / 74.9-76.7	89.6-90.7 / 75.2-78.2	99.9-100 / 99.7-100	80.7-82.5 / 78.2-80.1	78.5 / 70.2	-
P3	77.9-95.2 / 63.8-97.3	89.4-100 / 93.3-100	91.9-95.1 / 95.0-97.3	90.4-91.4 / 94.5-95.3	83.0-84.0 / 88.7-89.3	-	-	-	-
C1	90.1-91.5 / 87.1-88.5	98.0-100 / 96.8-100	96.4 / 94.6	91.3-96.4 / 83.1-94.1	96.0-96.5 / 94.1-95.2	93.3 / 90.7	92.4-93.2 / 89.6-92.0	81.4-81.6 / 72.8-74.7	85.3 / 77.4 (AF014811)
Nlb	91.1-93.6 / 91.8-98.0	93.4-94.0 / 98.0-98.4	92.1-94.1 / 97.8-98.4	100 / 100	94.5-94.6 / 98.6-98.8	90.5-92.0 / 96.9-98.6	91.0-91.5 / 96.9-97.4	82.9-84.0 / 94.9-95.9	-
Coat Protein	82.7-89.4 / 89.2-93.5	91.9-92.1 / 94.6-94.9	91.7 / 94.9	92.4-100 / 95.6-100	91.8-94.8 / 96.4-97.1	92.1-94.9 / 94.6-95.6	94.6-95.1 / 94.2-95.6	92.8-94.8 / 95.3-96.4	84.8-87.3 / 90.3-91.0

Table 2 Recombination sites detected in ZYMV genomes using seven non-parameteric methods showing the parental and recombinant sequences

Recombinant Sequence	Recombinant genome position		Gene/region	*Minor Parental Sequence(s)	*Major Parental Sequence(s)	**Detection methods
	Begin	End				
AB369279.1-RDA (South Korea)	82	2624	Partial 5' UTR, complete P1, HC-Pro and partial P3 gene	MK956829.1-isolate Z-104 (Italy)	Unknown	1, 3, 4, 5, 6, 7
AB188115.1-isolate: Z5-1 (Japan)	215	890	Partial P1 gene	Unknown	MF072714.1-isolate ZYMV-Trini3 (Trinidad and Tobago)	1, 2, 4, 5, 6, 7
MF072714.1-isolate ZYMV-Trini3 (Trinidad and Tobago)	226	914	Partial P1 gene	Unknown	MG967620.1- isolate SB-02 (India)	4, 5, 6
KX499498.1-isolate Vera (Spain)	284	1238	Partial P1 and HC-Pro gene	MF684760.1-isolate Kurdistan (Iran)	AF127929.2- isolate TW-TN3 (Taiwan)	1, 2, 3, 4, 5, 6, 7
MK956829.1-isolate Z-104 (Italy)	295	626	Partial P1 gene	Unknown	MF072714.1- isolate ZYMV-Trini3 (Trinidad and Tobago)	1, 2, 3, 4, 5, 6, 7
KX664482.1-ZYMV-WS (China)	888	2648	Partial P1, complete HC-Pro and partial P3 gene	MF072714.1-isolate ZYMV-Trini3 (Trinidad and Tobago)	KF976713.1- isolate SE04T (Slovakia)	4, 5, 7
AB188116.1-isolate 2002 (Japan)	2624	3744	Partial P3, complete 6K1 and partial C1 gene	Unknown	MK956829.1-isolate Z-104 (Italy)	1, 2, 3, 4, 5, 6, 7
MK956829.1-isolate Z-104 (Italy)	2626	6490	Partial P3, complete 6K1, C1, 6K2, VPg and partial Nla gene	KF976712.1-isolate H (Czech Republic)	MF072714.1-isolate ZYMV-Trini3 (Trinidad and Tobago)	4, 5, 6
KX664482.1-ZYMV-WS (China)	7196	9354	Partail Nlb and partial Coat protein gene	AY188994.1-strain B (Israel)	MF072714.1-isolate ZYMV-Trini3 (Trinidad and Tobago)	4, 5, 6, 7
MF072712.1-isolate ZYMV-Trini1 (Trinidad and Tobago)	7257	2618	Partial P3, complete 6K1, C1, 6K2, VPg, Nla and partial Nlb gene	AJ515911-Isolate WM (China)	KF976713.1-isolate SE04T (Slovakia), EF062582-isolate NAT (Israel), EF062583-isolate AG (Israel)	1, 4, 5, 6, 7
KU528623.1-isolate IKA (Iran)	8482	8657	Partail Nlb and partial Coat protein gene	AF127929.2-isolate TW-TN3 (Taiwan)	Unknown	1, 2, 3, 6
KU528623.1-isolate IKA (Iran)	8658	8972	Partial Coat protein gene	Unknown	JN183062.1-isolate ZYMV-Fars (Iran)	1, 2, 3, 4, 5, 6, 7

*'Minor' and 'Major' parents refer to the parental isolates contributing the smaller and larger fractions of the recombinant's sequence, respectively.

**Detection methods RDP (1), GENECONV (2), BOOTSCAN (3), MAXIMUM CHISQUARE (4), CHIMERA (5), SISTER SCAN (6), 3SEQ (7)

Table 3 Relative gene transcript levels of ZYMV-CP in seedling samples after aphid and seed transmission.

Aphid transmission*	Ct mean (samples)	Ct mean (18S rRNA)	ΔCt	ΔΔCt	Relative gene expression level ($2^{-(\Delta\Delta Ct)}$)
Day 7	33.32 ± 1.94	16.32 ± 0.76	17.00	-5.57	47.50
Day 17	31.82 ± 0.89	16.57 ± 0.92	15.25	-7.32	159.79
Day 27	31.12 ± 1.86	16.89 ± 0.99	14.23	-8.34	324.03
Day 37	28.62 ± 0.92	16.79 ± 0.69	11.83	-10.74	1710.26
Day 47	23.72 ± 0.77	16.57 ± 0.86	7.15	-15.42	43841.21
Day 57	21.45 ± 0.67	16.88 ± 0.71	4.57	-18.00	262144.00
Seed transmission**					
Day 7	34.75 ± 0.92	16.88 ± 0.87	17.87	-4.70	25.99
Day 17	32.25 ± 1.57	16.79 ± 0.72	15.46	-7.11	138.14
Day 27	29.5 ± 1.02	16.94 ± 0.91	12.56	-10.01	1031.12
Day 37	27.5 ± 0.98	16.08 ± 0.89	11.42	-11.15	2272.40
Day 47	24.15 ± 0.86	16.29 ± 0.84	7.86	-14.71	26801.01
Day 57	19.32 ± 1.42	16.56 ± 0.55	2.76	-19.81	919187.72
Positive control	20.78 ± 0.37	16.69 ± 0.42	4.09	-18.48	365623.68
Negative control	39.46 ± 0.29	16.39 ± 0.38	22.57	0.00	1.00

Ct mean; *n = 30, **n=6 (Includes three replicates of each sample), 18S rRNA; Endogenous control.

Supplementary Table 1 List of primers used for diagnosis and amplification of ZYMV partial genome fragments

S. No.	Primer name	Primer sequence (5' - 3')	Target Size	Target fragment (F) in genome
1	ZF1	AATCAACGAACAAGCAGACGA	971	F1; Partial 5' UTR and partial P1 gene
2	ZR1	CCTCGAATCACCAAATGTGGC		
3	ZF2	TGTAGTGGGTGGGTGTTAGGCA	1093	F2; Partial P1 gene and partial HC-Pro gene
4	ZR2	GCTCCTATAACTAGGTGCCTCTTCGT		
5	ZF3	GTGTCTCCAAACAGAATGG	1189	F3; Partial HC-Pro gene and partial P3 gene
6	ZR3	GCTCTTCGCATGTACTCGGG		
7	ZF4	GCGATTGAAGCAATTCTCGAT	1043	F4; Partial P3, complete 6K1 and partial C1 gene
8	ZR4	GGATCGCCTGCCAACTGTCTAC		
9	ZF5	CAACGAGCTTGCCCGCACATCTTGCCAA	1142	F5; Partial C1 gene
10	ZR5	GCACTCCAATGCGTTCATACTCAC		
11	ZF6	ATGCACCCATTGATACACGAAG	1161	F6; Partial C1 gene, complete 6K2 gene and partial VPg
12	ZR6	GTGCTTCTTCAGTGCCCTTGCCC		
13	ZF7	CGTTAGTGCAGGAGGAGTTTGG	1060	F7; Partial VPg gene, complete Nia gene and partial NIb gene
14	ZR7	GTTCCCTTCACAGCTTTCGTAGACC		
15	ZF8	GATGAACAACCAGGGCCTGAAT	1101	F8; Partial NIb gene
16	ZR8	TCTCAATCTCCTCGCAGTTCCAACC		
17	ZF9	GTGGGCAACCCTCAACGGTGGTGG	1176	F9; Partial NIb gene and partial Coat protein gene
18	ZR9	CTCATTTCTATGTATGCCTCCGC		
19	ZF10	AAAATGCAAAGCCAACGCTGCG	571	F10; Partial Coatprotein gene and 3' UTR
20	ZR10	AGGCTTGCAAACGGAGTCTAATCTCG		
21	CP-fwd	GCTCCATACATAGCTGAGAC	1100	Partial NIb gene and complete Coat protein gene and 3' UTR
22	CP-rev	AACGGAGTCTAATCTCGAGC		
23	ZYMVRT - F1	GAGAAATGCAGAGGCACCATACATGCCG	183	Partial Coat protein gene
24	ZYMVRT - R1	GGCCAAACAACCTTGAAGAAACATTGC		

Supplementary Table 2 List of isolates used in phylogenetic and recombination analyses in the study

S. No	GenBank Acc No	Isolate name	Crop	Country	S. No	GenBank Acc No	Isolate name	Crop	Country
1	KT598222	10itSDE	*Cucurbita maxima*	Argentina	33	MF072713	ZYMV-Trini2	Pumpkin	Trinidad and Tobago
2	MN598574	Knx-25	*Citrullus lanatus*	Australia	34	MF072714	ZYMV-Trini3	Pumpkin	Trinidad and Tobago
3	MN598577	Per-1	*Cucurbita pepo*	Australia	35	MF072715	ZYMV-Trini4	Pumpkin	Trinidad and Tobago
4	MN598579	Qld-5	*Cucumis melo*	Australia	36	KJ875864	USA	*Cucurbita pepo*	USA
5	MN364667	Brazil	*Watermelon cv. Manchester*	Brazil	37	JN192428	1st	*Cucurbita pepo*	USA
6	AJ307036	CU	Cucumber	China	38	JN192426	3rd	*Cucurbita pepo*	USA
7	AJ515911	WM	Watermelon	China	39	JN192424	5th	*Cucurbita pepo*	USA
8	AJ316229	WG	*Benincasa hispida*	China	40	JN192422	7th	*Cucurbita pepo*	USA
9	KF976712	H	*Cucurbita pepo*	Czech Republic	41	JN192419	E71	*Cucurbita pepo*	USA
10	MG967620	SB02	Soybean	India	42	JN192420	E72	*Cucurbita pepo*	USA
11	JN183062	ZYMV-Fars	*Cucurbita pepo*	Iran	43	JN192412	E81	*Cucurbita pepo*	USA
12	KU198853	SANRU	*Cucurbita pepo*	Iran	44	JN192413	E82	*Cucurbita pepo*	USA
13	KU528623	IKA	Squash	Iran	45	JN192409	F71	*Cucurbita pepo*	USA
14	EF062582	NAT	Cucurbits	Israel	46	JN192410	F72	*Cucurbita pepo*	USA
15	EF062583	AG	Cucurbits	Israel	47	JN192411	F73	*Cucurbita pepo*	USA
16	MK956829	Z-104	*Cucurbita pepo*	Italy	48	JN192406	F82	*Cucurbita pepo*	USA
17	AB188115	Z5-1	Cucumber	Japan	49	JN192407	F83	*Cucurbita pepo*	USA
18	AB188116	2002	Cucumber	Japan	50	JN192408	F84	*Cucurbita pepo*	USA
19	MH700748	15B	Cucumber	Papua New Guines	51	KC665630	FG2	*Cucurbita pepo*	USA
20	MH700750	16B	Cucumber	Papua New	52	KC665628	FG4	*Cucurbita pepo*	USA

				Guines					
21	L29569	Reunion Island	-	Reunion Island	53	JN192414	G61	*Cucurbita pepo*	USA
22	AF014811	Singapore	-	Singapore	54	JN192416	G71	*Cucurbita pepo*	USA
23	DQ124239	Kuchyna	*Cucurbita pepo*	Slovakia	55	JN192417	G72	*Cucurbita pepo*	USA
24	KF976713	SE04T	*Cucurbita pepo*	Slovakia	56	JN192418	G73	*Cucurbita pepo*	USA
25	AB369279	RDA	*Cucurbita pepo*	South Korea	57	KJ923767	leaf1	*Cucurbita pepo*	USA
26	AY278998	KR-PA	*Cucurbita moschata*	South Korea	58	KJ923769	leaf23	*Cucurbita pepo*	USA
27	AY278999	KR-PE	*Cucurbita moschata*	South Korea	59	KC665633	SG3	*Cucurbita pepo*	USA
28	AY279000	KR-PS	*Cucurbita moschata*	South Korea	60	KC665634	SG4	*Cucurbita pepo*	USA
29	KX499498	Vera	*Cucurbita pepo*	Spain	61	KC665635	SG5	*Cucurbita pepo*	USA
30	AF127929	TW-TN3	*Luffa cylindrica Roem.*	Taiwan	62	L31350	ZYMPRO POLR	*Cucurbita pepo*	USA
31	AM422386	begonia	Begonia	Taiwan	63	JQ716413	ZYMV_ PA_2006	*Cucurbita pepo*	USA
32	MF072712	ZYMV-Trini1	Pumpkin	Trinidad and Tobago					

Supplementary Table 3 List of variable amino acids in the polyprotein region of ZYMV between ZYMV-Trini isolates and their close relative isolates NAT (Israel), AG (Israel), SE04T (Slovakia)

Amino acid position	Variable amino acids*	Amino acid position	Variable amino acids*	Amino acid position	Variable amino acids*	Amino acid position	Variable amino acids*
4	I/./V/V/V/V	272	H/./P/P/P/P	960	E/./K/K/K/K/K	2580	A/./T/T/T/T
14	A/./T/././.	276	K/./G/G/G/G	965	A/./T/././.	2658	S/./C/C/C/C
15	K/./Q/Q/Q/Q	283	H/./Y/Y/Y/Y	980	V/./A/A/A/A	2745	K/./R/R/R/R
16	T/./P/P/P/P	333	L/./P/P/P/P/P	981	H/./Y/Y/Y/Y	2775	D/./E/E/E/E
17	E/./A/A/A/A	352	I/./T/T/T/T	1037	N/./K/K/K/K	2803	G/./D/D/D/D
19	C/./Y/Y/Y/Y	396	D/./N/././.	1038	T/./N/N/N/N	2806	P/./T/T/T/T
24	V/./A/A/A/A	407	I/./V/V/V/V/V	1119	V/./I/I/I/I/I	2811	T/./A/A/A/A/A
41	L/./P/P/P/P	414	A/./S/S/S/S	1127	V/./I/I/I/I	2817	D/./N/././.
47	M/./T/T/T/T	416	Q/./R/R/R/R	1171	S/./N/N/N/N	2818	K/./N/././.
68	H/./N/N/N/N/N	431	V/./A/././.	1381	F/./Y/Y/Y/Y	2828	V/./A/././.
96	V/./I/././.	470	M/./L/L/L/L	1394	Q/./H/H/H/H	2834	S/./G/G/G/G/G
101	S/./N/N/N/N	490	I/R/R/R/R/R/R	1652	S/./G/G/G/G	2838	V/./M/././.
165	T/./K/N/K/K	503	L/./F/././.	1662	I/./V/V/V/V	2839	A/./V/V/V/V
167	A/./V/V/V/V	527	L/./././V/.	1682	V/./I/././.	2841	V/./A/A/A/A
171	I/./T/T/T/T	545	V/./I/I/I/I	1717	A/./G/G/G/G	2842	T/./K/K/K/K
173	Q/./L/L/L/L	572	T/./I/././.	1745	D/./N/N/N/N	2917	E/./D/D/D/D
196	I/./M/M/M/M	578	S/./N/././.	1793	R/./K/K/K/K	2945	V/./F/F/F/F
214	C/./Y/Y/Y/Y	589	V/./I/I/I/I	1796	M/./V/V/V/V/V	2951	Q/./E/E/E/E/E
252	E/./K/././.	638	A/./T/././.	1816	R/./Q/Q/Q/Q	2958	K/./I/I/I/I
253	Q/./R/R/R/R/R	679	V/./I/I/I/I	1953	S/./N/N/N/N	2959	P/./L/L/L/L
255	C/./R/R/R/R	756	M/./L/L/L/L/L	2073	M/./T/T/T/T	3053	T/./I/I/I/I
264	S/./G/G/G/G	768	T/./S/S/S/S	2381	F/./V/V/V/V/V		
265	G/./R/R/R/R	788	L/./I/I/I/I	2395	F/./Y/Y/Y/Y		
267	V/./G/G/G/G	825	V/./I/I/I/I	2516	D/./N/././.		

*Isolates AG/NAT/SE04T/ZYMV-Trini1/ZYMV-Trini2/ZYMV-Trini3/ZYMV-Trini4. '.' representing identical amino acids.

Supplementary Table 4 Amino acid composition in the polyprotein region between ZYMV-Trini isolates and their close relative isolates NAT (Israel), AG (Israel), SE04T (Slovakia)

Amino Acid	Numbers*	Mol %
Ala A	189/189/189/189/189/189/189	6.13/6.13/6.13/6.13/6.13/6.13/6.13
Cys C	54/54/54/52/52/52/52	1.75/1.75/1.75/1.69/1.69/1.69/1.69
Asp D	160/160/157/160/160/160/160	5.19/5.19/5.1/5.19/5.19/5.19/5.19
Glu E	219/219/218/218/218/218/218	7.11/7.11/7.08/7.08/7.08/7.08/7.08
Phe F	156/156/156/154/154/154/154	5.06/5.06/5.06/5/5/5/5
Gly G	179/179/180/183/183/183/183	5.81/5.81/5.84/5.94/5.94/5.94/5.94
His H	90/90/89/87/87/87/87	2.92/2.92/2.89/2.82/2.82/2.82/2.82
Ile I	173/172/175/175/175/175/175	5.62/5.58/5.68/5.68/5.68/5.68/5.68
Lys K	219/219/220/220/219/220/220	7.11/7.11/7.14/7.14/7.11/7.14/7.14
Leu L	272/272/271/273/273/272/273	8.83/8.83/8.8/8.86/8.86/8.83/8.86
Met M	93/93/92/89/89/89/89	3.02/3.02/2.99/2.89/2.89/2.89/2.89
Asn N	154/154/160/159/160/159/159	5/5/5.19/5.16/5.19/5.16/5.16
Pro P	103/103/104/105/105/105/105	3.34/3.34/3.38/3.41/3.41/3.41/3.41
Gln Q	117/117/115/114/114/114/114	3.8/3.8/3.73/3.7/3.7/3.7/3.7
Arg R	168/169/170/172/172/172/172	5.45/5.49/5.52/5.58/5.58/5.58/5.58
Ser S	197/197/195/192/192/192/192	6.39/6.39/6.33/6.23/6.23/6.23/6.23
Thr T	194/194/195/193/193/193/193	6.3/6.3/6.33/6.26/6.26/6.26/6.26
Val V	211/211/208/207/207/208/207	6.85/6.85/6.75/6.72/6.72/6.75/6.72
Trp W	38/38/38/38/38/38/38	1.23/1.23/1.23/1.23/1.23/1.23/1.23
Tyr Y	94/94/94/100/100/100/100	3.05/3.05/3.05/3.25/3.25/3.25/3.25

*Isolates AG/NAT/SE04T/ZYMV-Trini1/ZYMV-Trini2/ZYMV-Trini3/ZYMV-Trini4.

Figures

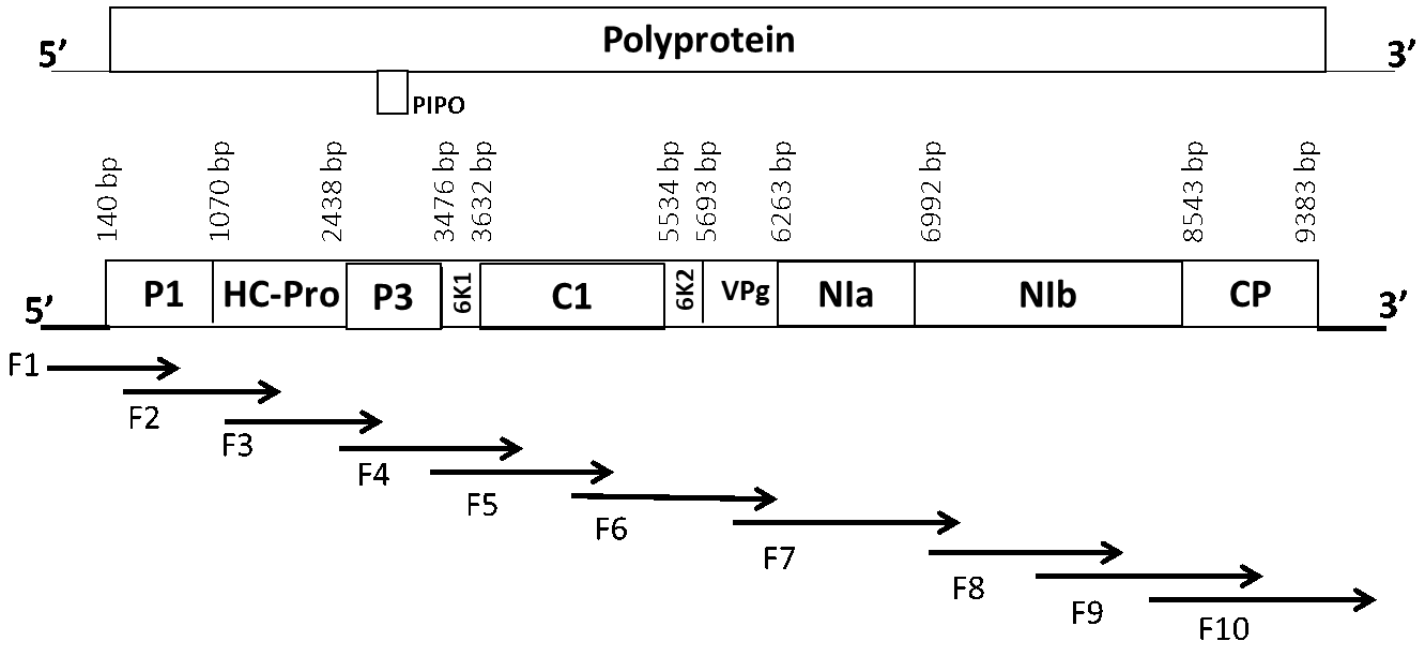

Figure 1

Position of fragments (F1 – F10) amplified in PCR using a list of primers designed (Supplementary Table 1) to derive the complete genome of ZYMV-Trini isolates.

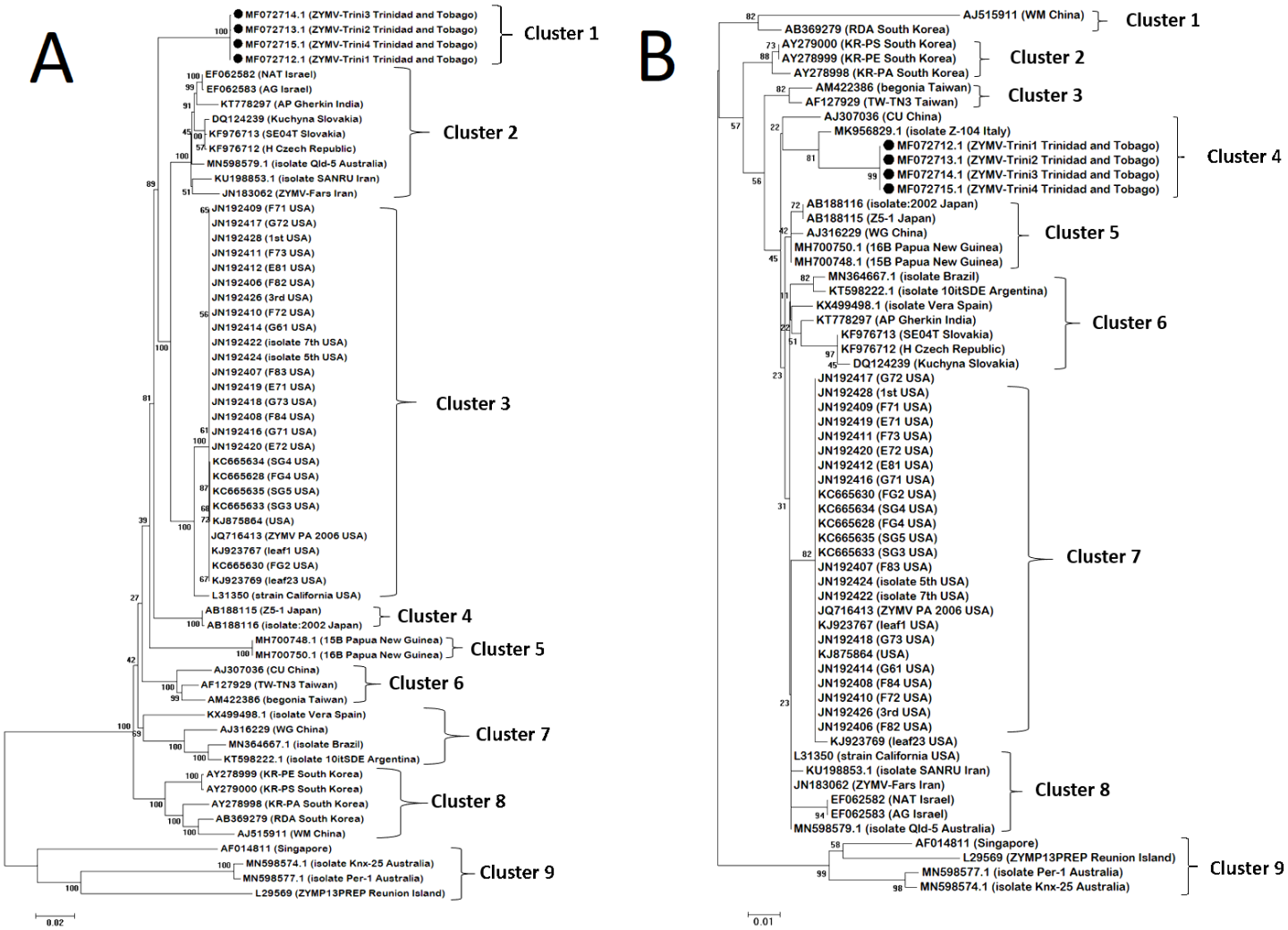

Figure 2

Maximum Likelihood phylogram constructed using the Tamura-Nei model in Mega X software with 1000 bootstrap replications. a) Phylogenetic tree constructed based on full length nt sequences of ZYMV isolates. b) Phylogenetic tree constructed for CP amino acid sequences of ZYMV isolates. Branch lengths are proportional to the number of substitution. ● (ZYMV-Trini isolates from this study).

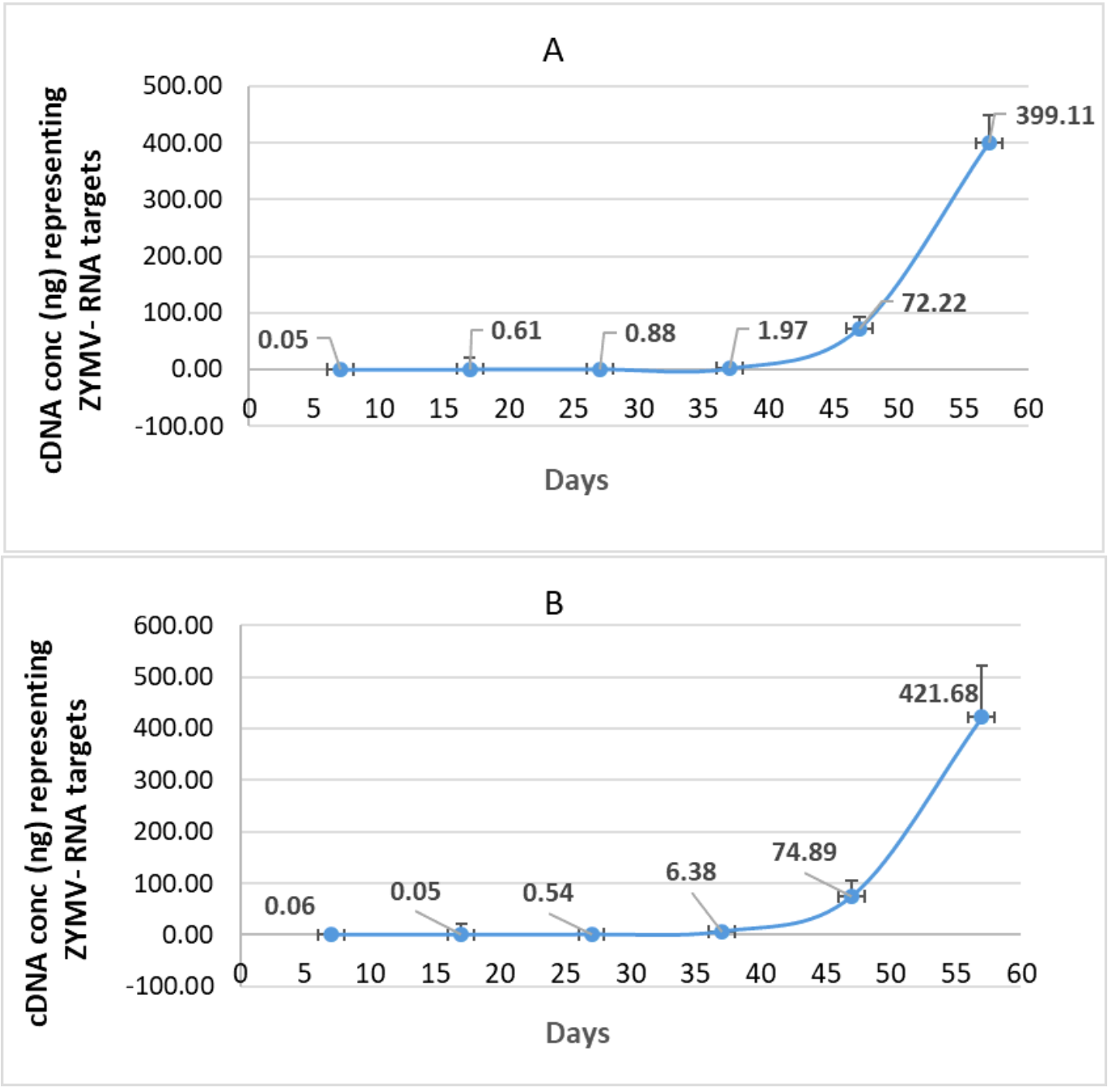

Figure 3

Quantification of ZYMV-RNA targets using RT-qPCR. A: leaf samples collected after aphid transmission. B: leaf samples collected after germination of seeds from the fruits of ZYMV positive plants. Progression of ZYMV-RNA targets was monitored using qPCR with specific primers and correlating CT values to cDNA quantity generated from standard curves. Data points represents means (A (n = 30), B (n=6) ± SD. Means were analysed by repeated measures ANOVA at 5% confidence levels.

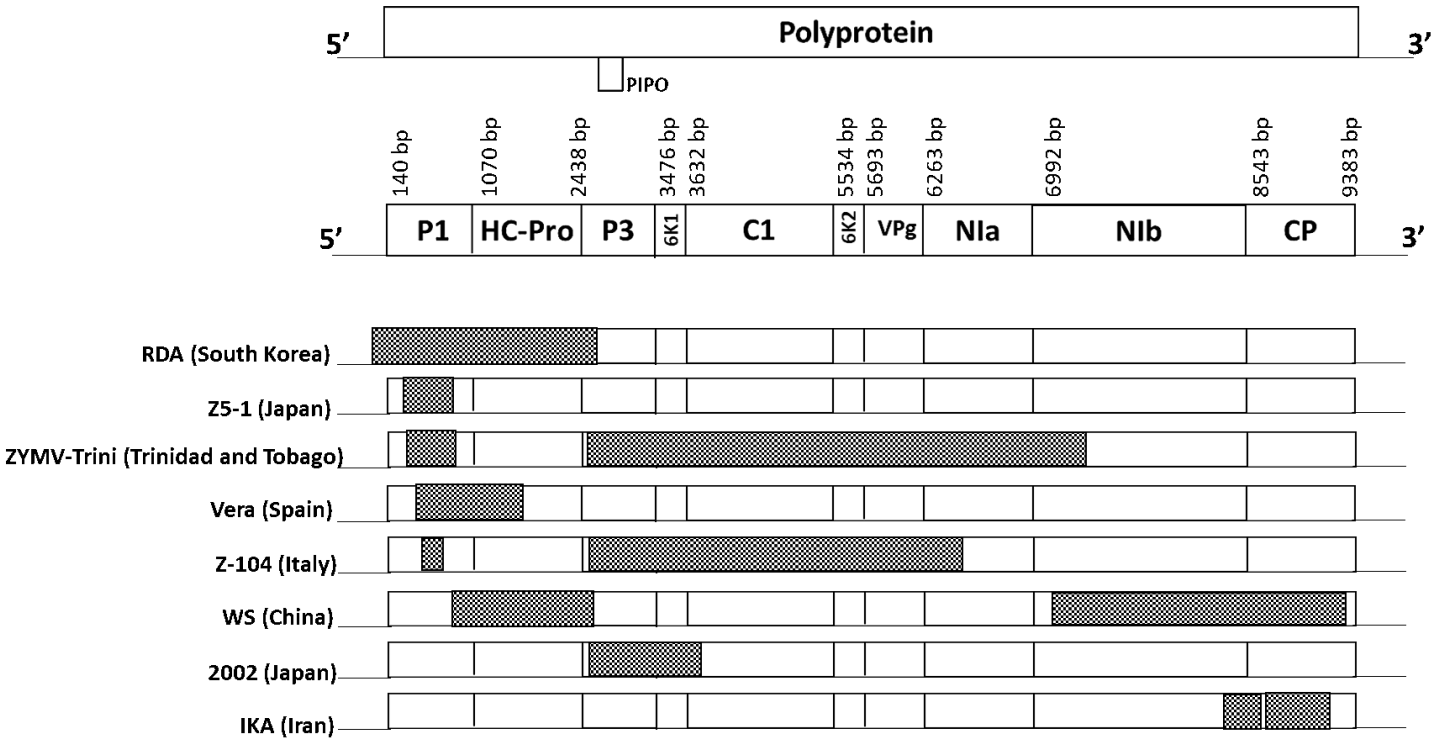

Figure 4

Recombinant sites detected in the genome of ZYMV. At least twelve recombination events were found in nine isolates from seven different countries. The shaded regions are putative recombinant sites assessed by eight recombination detection methods in RDP5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure1.png