Genome characterization of zucchini yellow mosaic virus infecting cucurbits reveals the presence of a new genotype in Trinidad and Tobago in the Caribbean region

Chinnaraja Chinnadurai1 · Mounika Kollam1 · Adesh Ramsubhag1 · Jayaraj Jayaraman1

Received: 5 December 2020 / Accepted: 8 February 2021 / Published online: 3 April 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract
Zucchini yellow mosaic virus (ZYMV) is a member of the genus Potyvirus that is becoming a serious pathogen of pumpkin and other cucurbits in Trinidad and Tobago and the entire Caribbean region. In this study, four ZYMV isolates infecting pumpkin in Trinidad and Tobago were characterized by complete genome sequencing. Phylogenetic analysis showed 5.9–6.0% nt and 7.7–7.9% aa sequence divergence in comparison to the most closely related isolates NAT and AG from Israel and SE04T from Slovakia. Based on the variations in the complete genome sequence as well as individual gene sequences, a new genotype, designated ZYMV-Trini, is proposed for these isolates. Among the gene sequences of ZYMV-Trini isolates, the greatest variation was observed in the HC-Pro gene, with 20.8% aa sequence divergence from their closest relatives, whereas the least variation was observed in the NIb, P3, and CP genes, with 1.8–2.2% aa sequence divergence. This study also showed that transmission of ZYMV can occur through seeds, but this was less common than transmission via the aphid 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 recombination analysis, it could be concluded that the isolates SE04T from Slovakia, NAT from Israel, and AG from Israel have made major contributions to the genome architecture of ZYMV-Trini isolates.

Introduction
Zucchini yellow mosaic virus (ZYMV) is a member of the genus Potyvirus within the family Potyviridae. ZYMV was first reported in Italy in 1973 [27], and it subsequently spread worldwide, causing devastating epidemics in tropical, subtropical, and temperate regions. Members of the plant family Cucurbitaceae are the primary hosts of ZYMV, and disease symptoms include severe mosaic, yellowing, distortion of leaves, stunting of plant growth, severe fruit deformation, and color cracking [8]. The symptoms can render fruits unmarketable and cause yield losses 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 to be associated with transmission in the field [18].

The ZYMV genome is a ~9.6-kb positive-sense single-stranded RNA molecule. The genome has one open reading frame (ORF) encoding a single polyprotein precursor that is subsequently processed by three virally encoded proteases to produce 10 functional small mature proteins: P1 (protease), HC-Pro (helper component/protease), P3, 6K1, CI (cylindrical inclusion protein), 6K2, NIa (nuclear inclusion protein a), VPg (genome-linked viral protein), NIb (nuclear inclusion protein 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 5' untranslated region (UTR) of ZYMV contains 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/...
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 the highest incidence of ZYMV infection (74%) in the dry season. Furthermore, coinfection with ZYMV and squash mosaic virus (SqMV) in cucurbits in Trinidad and Tobago has also been reported [3].

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

Materials and methods

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 Grove on the island of Trinidad [3] were used for complete viral genome sequencing. Total RNA was extracted from leaf samples (1 g) using TRI Reagent (Sigma, USA) following the manufacturer’s protocol.

RT-PCR and sequencing

Diagnosis of ZYMV was carried out by PCR using the primers CP-forward (5’-GCTCCATACATAGCTGAGAC-3’) and CP-reverse (5’-AACGGAGTCTAATCTCGAGC-3’), targeting a portion of the coat protein gene (1100-nt) of ZYMV. Ten different pairs of primers targeting overlapping fragments of the ZYMV ORF were designed (Fig. 1; Supplementary Table 1) to determine the complete genome sequences of the Trinidad isolates. An ImProm-II™ Reverse Transcription System (Promega, USA) was used for the synthesis of complementary DNA using 1 μg of RNA. All PCR reactions were performed in a thermocycler (Techne, USA). Each PCR reaction contained 100 ng of cDNA, 1 unit of Pfu DNA polymerase, 10X buffer with MgSO4, 0.5 μl of 10 mM dNTP mix, 50 pmol of primer pairs, and sterile Milli-Q Water to a final volume of 25 μl. The PCR conditions were as follows: initial denaturation at 94 °C for 4 min, followed by 30 cycles of 94 °C for 1 min, 54-60 °C for 1 min, and 72 °C for 1 min, and then a final primer extension step for 10 min at 72 °C. Amplicons were visualized by electrophoresis on a 1.5 % agarose gel stained with ethidium bromide. Amplicons from three replicates of each PCR were purified from the gel using a Gen-Elute Gel Extraction Kit (Sigma, USA), cloned into pGEM-T Vector (Promega, USA), and sequenced in both strands by the Sanger method at Macrogen Inc, USA. The complete genome sequences of the ZYMV isolates from Trinidad were constructed by aligning all partially overlapping fragments of ZYMV ORF sequences with reference genome sequences of ZYMV from the NCBI GenBank database using the bioinformatics software MEGA X [20].

Phylogenetic analysis

All of the nucleotide (nt) and amino acid (aa) sequence fragments corresponding to the polyprotein of ZYMV were aligned separately with reference sequences obtained from the GenBank database (Supplementary Table 1) 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 the RefSeq database. Phylogenetic analysis was carried out using the bioinformatics software MEGA X [20].
out using the complete genome sequences of the ZYMV Trinidad isolates and 63 reference sequences obtained from GenBank representing different geographical regions of the world (Supplementary Table 2). Maximum-likelihood trees were generated in MEGA X software using the Tamura-Nei model for the complete genome nt sequence and the aa sequences of the individual proteins P1, HC-Pro, P3, CI, N Ib, and CP with 1000 bootstrap replications. Similarity matrices showing the percentage nt and aa sequence identity for all of the clusters in the phylograms were also generated in MEGA X [20].

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 (confirmed by PCR) for acquisition feeding for 48 h. The viruliferous aphids were then transferred individually to 15 sterile pumpkin seedlings for 48 h for inoculation feeding in a netted greenhouse box. After inoculation, the aphids were killed using malathion treatment. After seven days, total RNA was extracted from leaf samples (1 g of third leaf) from inoculated and control seedlings using TRI Reagent (Sigma, USA). The RNA was reverse transcribed, and PCR amplification of the cDNA was 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 their identity by BLAST (NCBI), which confirmed the infection by ZYMV. Leaf samples from 10 of the 12 ZYMV-positive seedlings were collected every 10 days until flowering, and RNA was extracted. The virus titres in these 10 samples were determined using three replicates by RT-qPCR to assess the progression of infection at different growth stages.

In addition, 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 four times with distilled water to ensure removal of all contaminants. All of the surface-sterilized seeds were pooled together, and 100 seeds were collected randomly and planted in individual pots. Leaf samples were collected 7 days after germination. PCR reactions confirmed ZYMV infection in 2 out of 100 seedlings raised from the seeds. Leaf samples from those two positive seedlings were collected every 10 days thereafter for virus quantification as before.

For cDNA synthesis, RNA (500 ng) was reverse transcribed using MultiScribe Reverse Transcriptase (Invitrogen, USA), primed with 40 nmol of primer ZYMVRT-R1 (5’-GGCCAAAACACTTGAAGAACATTGC-3’) in a 20-µL reaction following the manufacturer’s protocol, using a thermocycler (Techne, USA). Real-time quantitative PCR was performed with three replicates of each sample with 500 ng of cDNA template, 12.5 µL of SYBR Green JumpStart™ Taq ReadyMix™ (Sigma, USA), and 50 nM primers ZYMVRT-F1 (5’-GAGAAATGCAGAGGCACC ATACATGCCG-3’) and ZYMVRT-R1 (5’-GGCCAAAAC ACTTGAAGAACATTGC-3’), targeting a 181-nt region of the ZYMV coat protein gene. RT-qPCR (25 µL) 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 of the samples. The optimised RT-qPCR conditions were 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Melt curve analysis was performed at 60 °C for 15 s to ensure that a homogenous amplification product was produced. ZYMV-infected samples from greenhouse pots were used as positive controls.

Relative standard curve analysis was done using 500 ng of cDNA from the ZYMV-positive control, and a tenfold serial dilution was made to generate the standard curve. The threshold cycle number (Ct) of each dilution was plotted against standard concentrations of cDNA, and a standard curve was constructed. A regression line generated for the standard curve was used to determine the titre of ZYMV in the individual samples at each sampling time.

In addition, the expression level of the ZYMV-CP target was assessed by the relative quantification method using 2^(-ΔΔCt) values. The ΔCt value was determined by subtracting the Ct value of the endogenous control (18S rRNA) from that of the test sample (ZYMV-CP). The ΔΔCt value was determined by subtracting the ΔCt value of the negative control from that of the sample. PCR-confirmed greenhouse-grown pumpkin plants were used as positive and negative controls for qPCR analysis.

Recombination analysis

Recombination analysis was performed using all 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 cutoff of 0.05 and default settings were used throughout the analysis, and only events detectable by
three or more different methods were subjected to further analysis.

**Results**

RT-PCR analysis with the diagnostic primers CP-for and CP-rev confirmed ZYMV infection in all of the samples collected from Barrackpore, Macoya, Las Lomas, and Orange Grove. Portions of the ZYMV ORF 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 database (Supplementary Table 2), and the complete genome sequence was constructed for four Trinidad isolates, which were designated as “ZYMV-Trini” isolates. The complete genome sequence was aligned with multiple reference isolates collected from the GenBank database (Supplementary Table 2), and the complete genome sequence was deposited in the GenBank database under the following accession numbers: ZYMV-Trini1, MF072712; ZYMV-Trini2, MF072713; ZYMV-Trini3, MF072714; ZYMV-Trini4, MF072715.

Pairwise comparisons of the nt sequences of the ZYMV-Trini isolates showed that they were 99.9-100 % identical (Table 1). A phylogram constructed based on the complete genome nt sequences of these isolates and 63 reference sequences from various countries showed that the ZYMV-Trini isolates formed a separate cluster (Fig. 2). The ZYMV-Trini isolates were most closely related (94.0-94.1 % nt sequence identity) to the reference isolates NAT (Israel), AG (Israel), and SE04T (Slovakia). The polyprotein of the ZYMV-Trini isolates showed 93 variable amino acids in comparison with NAT (Israel), AG (Israel), and SE04T (Slovakia) (Supplementary Tables 3 and 4). The isolates Per-1 (Australia), Knx-25 (Australia), and ZYMPP13PREP (Reunion Island) were found to have the least aa sequence similarity and 23.2-23.6 % nt sequence divergence from the ZYMV-Trini isolates (Table 1).

Phylogenetic analysis based on aa sequences of individual proteins of the ZYMV-Trini and reference isolates showed that the four local isolates formed a separate and distinct cluster for HC-Pro, CI, and NB (Supplementary Fig. 1). Pairwise comparison with all of the P1 reference gene sequences showed that the Trini isolates were closely related to isolate TW-TN3 (Taiwan) with 97.6 to 97.7 % nt sequence identity. Similar analyses for the HC-Pro and NB genes showed that ZS-1 (Japan) was the most closely related isolate to the ZYMV-Trini isolates, with 90.7 % and 94.6 % nt sequence identity, respectively. Pairwise comparisons with reference sequences for the P3, CI, and CP genes showed that the Trini isolates were closely related to isolate Z-104 (Italy), with 97.6, 98.0, and 97.2 % nt sequence identity, respectively (Table 1, Fig. 2, Supplementary Fig. 1).

The isolate ZYMPP13PREP (Reunion Island) was found to have the lowest nt sequence similarity to the Trini isolates in the regions encoding CI (81.4 %) and NB (82.9 % identity), whereas the isolate WM (China) showed the highest nt sequence similarity in the CP gene (82.7 % identity), and the isolate Singapore (Singapore) showed the least similarity in the regions encoding P1 (62.8-62.9 % identity) and HC-Pro (80.7 % identity) (Table 1).

### Table 1 Nucleotide (nt) and amino acid (aa) sequence identity (%) of ZYMV-Trini isolates to isolates from 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 / 91.2-92.8 | 90.8-91.9 / 91.2-91.7 | 89.4-90.8 / 88.7-89.4 | 93.6-93.7 / 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 | 88.9-87.9 / 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 / 78.2-80.1 | 80.7-82.5 / 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 / 77.3-79.0 |
| C1 | 90.1-91.5 / 98.0-100 | 98.0-100 / 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 / - | 81.4-81.6 / 72.8-74.7 | 85.3 / 77.4 | (AF014811) |
| Nb | 91.9-93.6 / 91.8-98.0 | 93.4-94.0 / 98.0-98.4 | 92.1-94.1 / 97.8-98.4 | 90.5-92.0 / 98.6-98.8 | 94.5-94.6 / 96.9-98.6 | 91.0-91.5 / 96.9-97.4 | 92.8-84.0 / 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 | 94.2-95.6 / 95.3-96.4 | 90.3-91.0 | - / - |
Progression of virus titre

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

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

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

Recombination analysis

Phylogenetic analysis of ZYMV based on complete genome nt sequences revealed a maximum of 23.6% variation among isolates worldwide. Recombination analysis using seven different detection methods detected
recombination sites throughout the genome of ZYMV. Twelve recombination sites were detected in eight different isolates from seven countries (Fig. 4).

In the ZYMV-Trini isolates, the first putative recombination event involved part of the region, the P3 complete 6K1, C1, 6K2, and VPg regions, part of the NIa region, with the isolates SE04T (Slovakia), NAT (Israel), AG (Israel) identified as a major parent and WM (China) as a minor parent. The second putative recombination event in the ZYMV-Trini involved part of the P1 region, and isolate SB-02 (India) was found to be a major parent. The hypothetical parental and daughter sequences that would fix the patterns of recombination sites were identified using seven non-parametric methods as implemented in RDP5 software. Isolates ZYMV-WS from China, ZYMV-Trini from Trinidad, and Z-104 from Italy were found to have two recombination sites in their genome. It is noteworthy that the ZYMV-Trini isolates were identified 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 (Supplementary Table 5).

Knowledge of ZYMV variability is essential for understanding the complexity of this virus and designing effective control strategies. Important biological, serological and molecular variability among ZYMV isolates has been described.

**Discussion**

ZYMV is becoming a serious pathogen in most cucurbit-growing regions of the world, with infection rates of at least 40% being reported in tropical and subtropical regions. Disease surveys in pumpkin have indicated that

---

Table 2 Relative gene transcript levels of ZYMV-CP in seedling samples after aphid and seed transmission

| Aphid transmission* | Mean Ct (samples) | Mean Ct (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                             |

* n = 30; ** n = 6 (includes three replicates of each sample); 18S rRNA, endogenous control
Zucchini yellow mosaic virus in Trinidad and Tobago

1 3

the ZYMV disease incidence can be as high as 75% in the 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 large variety of strains [9]. In this study, phylogenetic analysis revealed that the Trinidad isolates form a new genotype, ‘ZYMV-Trini’, since their complete nt sequences differ by 5.9-6.0% from those of their closest known relatives, including the isolates NAT and AG (Israel) and SE04T (Slovakia). The within-species genotype classification system uses a neutral nomenclature involving letters of the alphabet and Latinized numerals that avoid potentially misleading names [29, 30]. Phylograms of aa sequences also showed that the ZYMV-Trini isolates form a separate cluster based on HC-Pro, CI, and NIb aa sequences. In the phylogram based on P1 sequences, the isolate TW-TN3 (Taiwan) was found to be closely related to the Trini isolates, with 2.3–2.4% nt sequence divergence.

The capsid protein (CP) gene of potyviruses is widely used for typing isolates [36]. However, comparison of complete genome sequences allows a more comprehensive analysis of virus variability and may provide information about the evolutionary history of the isolate and the occurrence of major evolutionary events such as recombination, as has been demonstrated for various potyviruses [14, 39]. In the ZYMV-Trini isolates, NIb and CP were found to be highly conserved, whereas the HC-Pro gene had the most aa sequence variation when compared to 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 the ZYMV-Trini isolates from this study grouped together, suggesting that all of the ZYMV isolates from Trinidad and Tobago belong to the same genotype (data not shown). Among the other reference isolates, ZYM13PREP from Reunion Island differed the most from the Trini isolates, viz., 23.6% in the complete

... genome nt sequence, 18.6% in the CI aa sequence, and 7.1% in the NIb aa sequence.

Aphids are the most successful vectors of potyviruses due to certain features they possess [15], including their ability to deliver viral particles precisely via the stylet, their parthenogenetic mode of reproduction within a short span of time, the diversity of host plants they can infest, their ability to survive in adverse conditions, and their ability to disseminate over long distances [28, 33]. In a study in Greece, Katis et al. [18] found the most abundant aphid vectors of ZYMV to be *M. persicae*, *Aphis gossypii*, and *Aphis spiraecola*. We also reported *A. gossypii* to be a vector of ZYMV in Trinidad in our earlier preliminary study [3], and this was confirmed in the current study. Using RT-qPCR, we detected an incremental increase in ZYMV RNA in pumpkin seedlings following transmission by *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 protease (HC-Pro). Specifically, the DAG motif on the CP interacts with the PTK region of 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 the spread of ZYMV to cucurbit crops [5, 6, 23].

Generally, ~20% of viral plant pathogens are known to be seed-transmitted. The seed-to-seedling transmission rate of ZYMV was reported earlier to be low (1.6%) [17, 37]. A similar seed transmission rate (2%) was also observed in this study. Vertical transmission of ZYMV by seed is less common than horizontal transmission by aphids, and many studies have shown that the insect vector is an important factor influencing virus variation [35]. In both aphid and seed transmission experiments, the ZYMV titer increased steadily up to 37 days, but a rapid surge was observed by RT-qPCR analysis between 37 and 57 days postinfection.

Through recombination, viruses can gain pathogenicity or virulence or the ability to infect 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 it is in other RNA viruses. The detection of recombination sites using RDP5 software suggested that the entire ZYMV genome is prone to recombination, although hotspots are concentrated in the P1, HC-Pro, P3, and CP genes in several isolates. Recombination breakpoints in the ZYMV-Trini isolates were found in the P1 gene and between the P3 and Nlb genes. Maina et al. [29] reported the same pattern of recombination breakpoints in ZYMV populations in cucurbit crops in East Timor and northern Australian. Natural recombinants may emerge in virus populations only if they maintain relatively good fitness, which 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].

The isolates SE04T (Slovakia), AG (Israel), and NAT (Israel) were identified as the putative major parents of the ZYMV-Trini isolates. These parental isolates have 98.0-98.4% nt sequence identity to each other and 93.4-94.1% identity to the ZYMV-Trini isolates. Cucurbit cultivation in the Caribbean islands, including Trinidad and Tobago, is mainly dependent on imported seeds from countries such as Israel and China. ZYMV can also move to new locations in infected fruit from which aphids can acquire and spread the virus [24]. Introductions can also occur due to migrating birds carrying virus-infected seed in their intestines or to discarded infected cucurbit fruit being left behind by fishermen from neighboring 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 Australia or East Timor. The highest nucleotide sequence identity between a ZYMV sequence from PNG and elsewhere was 92.8%, and the authors suggested that this 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 note that ZYMV-Trini strains might have contributed genes through recombination to the isolates Z5-1 (Japan), Z-104 (Italy), and ZYMV-WS (China). However, more data need to be generated from the Caribbean island countries to study the genome dynamics of ZYMV-Trini strains and their genetic relationships to 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 important to understand the complexity of the variability of ZYMV isolates in order to establish effective field control measures.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00705-021-05048-4.

Acknowledgements The authors are grateful to ACP-EU (Grant no. PED/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. We thank 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, Perrin TM (1989) Effect of zucchini yellow mosaic virus on development and yield of cantaloupe (Cucumis-Melo). Plant Dis 73:317–320
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, Ramkisson A, Rajendran R, Tony ST, Ramesubhag A, 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 cucumber 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 9:267–276
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, Pitterero 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

Springer
15. Harris KF, Maramorosch K (1997) Aphids as virus vectors, 1st edn. Academic Press, Cambridge (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 Plam virus and the first definitive evidence of intraspecies recombination events. Phytopathology 106:407–416.

17. Johansen E, Edwards MC, Hampton RO (1994) Seed transmission of viruses: current perspectives. Annu Rev Phytopathol 32:363–386.

18. Kaisi NJ, Tsisipis JA, Lykourressis DP, Papapanayotou A, Margaritopoulos JT, Kokinis GM, Peridis 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 Phytopathol 154:293–302.

19. Khana V, Ali A (2019) Complete genome sequence of a Zucchini yellow mosaic virus isolated from pumpkin in Oklahoma. Microbiol Resour Announc 82:e01583-e1618.

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 X version 2.0. 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 (2008) Potential involvement of melon fruit in the long-distance dissemination of cucurbit potyviruses. Plant Dis 92:955–959.

25. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. https://doi.org/10.1093/ve/vev003.

26. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. https://doi.org/10.1093/ve/vev003.

27. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. https://doi.org/10.1093/ve/vev003.

28. Ng JC, Perry KL (2004) Transmission of plant viruses by aphid vectors. Mol Plant Pathol 5:505–511.

29. Niefelt 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 73:11900–11908.

30. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3:336–340.

31. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3:336–340.

32. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3:336–340.

33. Phypack EP, Shukla DD (1992) Coat protein phylogeny and systematics of potyviruses. Arch Virol Suppl 5:139–170.

34. Simmons HE, Holmes EC, Gildow FE, Bothe-Goralczyk MA, Stephenson AG (2011) Experimental verification of seed transmission of Zucchini yellow mosaic virus. Plant Dis 95:751–754.

35. Tcherepanov V, Elhers A, Upton C (2006) Genome annotation transfer utility (GATU): rapid annotation of viral genomes using a closely related reference genome. BMC Genomics 7:150. https://doi.org/10.1186/1471-2164-7-150.

36. 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 11:2099–2111.

37. 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 11:2099–2111.

38. 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 11:2099–2111.

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 11:2099–2111.

40. 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 11:2099–2111.