Journal of Cell and Molecular Research (2019) 10 (2), 76-89 DOI: 10.22067/jcmr.v10i2.76133

Research Article

Population Genetic Analysis of *Zucchini yellow mosaic virus* based on the CI Gene Sequence

Zohreh Moradi[1], Mohsen Mehrvar[1], Ehsan Nazifi[2*]

[1] *Department of Plant Pathology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran*
[2] *Department of Biology, Faculty of Basic Sciences, University of Mazandaran, Babolsar, Iran*

Received 22 October 2018 Accepted 27 January 2019

Abstract

Zucchini yellow mosaic virus (ZYMV) is one of the most economically important viruses infecting cucurbits worldwide. Population genetic analysis of ZYMV was conducted based on the virus cylindrical inclusion (CI) gene sequences of 10 isolates identified in this study and 94 other isolates from different countries in six continents: Asia, Europe, Oceania, Africa, North and South America. The overall mean value of nucleotide sequence diversity among all isolates was 0.074±0.006. Phylogenetic analysis showed that ZYMV isolates fell into three main phylogroups with significant F_{ST} values (>0.55) and almost tended to cluster according to their geographical position. Group I was predominant and contained isolates originated from different parts of the world. Iranian isolates clustered into group I, sharing 87.7-99.7% and 92.5–100% nucleotide and amino acid identity, with other isolates of this group. Group II was a new group that included only Singapore isolates. Group III including East Timor, Reunion Island and Australia-Kununurra isolates which were genetically differentiated from other populations. ZYMV populations from different geographic origins were composed of multiple lineages. With exception of the Oceanian population which was strongly differentiated from the American population, most other geographical populations showed low to moderate genetic differentiation. There was moderate to high level of gene flow despite large separating geographic distances. Analysis of the synonymous-to-nonsynonymous ratio showed strong purifying selection in the CI gene. The analyses indicated that in addition to selection, random processes such as genetic drift and founder effects are important determinants for the genetic structure of populations of ZYMV.

***Keywords*:** Cylindrical inclusion, Evolutionary forces, Genetic variability, *Zucchini yellow mosaic virus*

Introduction

Zucchini yellow mosaic virus (ZYMV; genus *Potyvirus,* family *Potyviridae*) is a damaging plant pathogen that infects a wide range of cucurbit crops worldwide (Desbiez and Lecoq, 1997; Lisa and Lecoq, 1984), with major economic impact and significant yield losses. The virus was first isolated in Italy in 1973, described in 1981 by Lisa et al. (1981), subsequently in France by Lecoq et al. (1981). As other potyviruses, ZYMV has flexuous filamentous particles of 680–730 nm long, which encapsidate a single-stranded, positive-sense RNA of approximately 10 kb. The viral genome, which has a poly (A) tail at its 3′end and a VPg structure at its 5′ end, consists of a unique large open reading frame (ORF) which encodes a single large polyprotein that is self-hydrolyzed after translation into 10 putative functional proteins (from N- to C-termini): P1 protein, helper component proteinase (HC-Pro), P3 protein, 6K1, cylindrical inclusion protein (CI), 6K2, nuclear inclusion protein a (NIa) (VPg+Pro), nuclear inclusion protein b (NIb) and coat protein (CP) (Adams et al., 2005a; Urcuqui-Inchima et al., 2001; Riechmann et al., 1992; Adams et al., 2012). In addition, a pretty interesting *Potyviridae* ORF, which is embedded in the P3-coding region, encodes a small putative protein PIPO (Chung et al., 2008). ZYMV, which is aphid-transmitted in a non-persistent manner (Lecoq et al., 1991; Desbiez et al., 1996; Gal-On, 2007), can infect wild and agronomically important cucurbit plants, some non-cucurbitaceous weeds, and some ornamental plants (Al-Musa, 1989; Desbiez and Lecoq, 1997; Coutts and Jones, 2005; Chen and Hong, 2008; Choi et al., 2002). Seed transmission, although at very low rates, has been reported in some cases (Desbiez and Lecoq, 1997; Tobias and Palkovics, 2003; Schrijnwerkers et al., 1991; Coutts et al., 2011; Simmons et al., 2011, 2013), which could explain ZYMV worldwide distribution (Desbiez et al., 2002). Several studies have been published in recent years on ZYMV biological and molecular variability in the world (Coutts et al., 2011; Desbiez et al., 1996, 2002; Glasa et al., 2007;

*Corresponding author E-mail: E.Nazifi@umz.ac.ir

Maina et al., 2017; Novakova et al., 2014; Yakoubi et al., 2008) as well as in Iran (Bananej et al., 2008; Masumi et al., 2011). Most of the molecular studies were based on analysis of CP and or partial NIb-CP sequences. Based on these phylogenetic analyses, ZYMV isolates have been classified into two or three major phylogroups (Desbiez et al., 2002; Zhao et al., 2003; Simmons et al., 2008; Ha et al., 2008a; Bananej et al., 2008; Yakoubi et al., 2008; Masumi et al., 2011; Coutts et al., 2011; Maina et al., 2017). On the other hand, in the absence of complete genomic sequence, cylindrical inclusion (CI)-coding region is the most suitable part for diagnostic and taxonomy purposes, rather than the CP (Ha et al., 2008b; Adams et al., 2005b, Lee et al., 1997). Molecular evolutionary studies of viruses focused on understanding effects of variation caused by mutation, recombination, selection pressure, and host or geography driven adaptation in viral populations (Moury et al., 2002; Gibbs and Ohshima, 2010). So, studying the molecular evolutionary history of plant viruses and understanding their genetic variation and the causative factors producing variation in viral populations is important for developing sustainable management strategies. Despite worldwide distribution of this virus, molecular evolution and population genetic structure are poorly understood and further investigation is required. This study was aimed to investigate population genetic structure and genetic diversity of ZYMV to identify the sources of genetic variation operating in the ZYMV population. It is based on analysis of the CI genomic region, which is a region that, to date, has not been analyzed in other studies. According to the Adams et al. (2005b) comparisons of the CI gene most accurately reflected those for the complete ORF, and this region would be the best for diagnostic and taxonomic studies if only a sub-portion of the genome were sequenced and was therefore selected for this study. Here, the CI nucleotide sequences of ten ZYMV isolates were obtained and analyzed together with those retrieved from the GenBank.

Materials and Methods

Virus sources, RT-PCR, cloning and sequencing

During the growing season of 2013, cucurbit and tomato plants with symptoms of ZYMV infection (including systemic mosaic, yellowing, vein clearing and banding, stunting, blistering, shoestring and leaf and fruit deformations) were collected from northern (Mazandaran, Golestan) and eastern (Razavi Khorasan) areas of Iran (Table 1). Total RNA was isolated and used as a template for reverse transcription (RT). One pair of degenerate primer including CI For/CI Rev corresponding to CI coding region in the potyvirus genome (Ha et al., 2008b) was used in the RT-PCR reactions. The first strand cDNA was synthesized using antisense primer and the Moloney murine leukemia virus (MMuLV) reverse transcriptase (Thermo Scientific, USA) according to the manufacturer's instructions. PCR was carried out using Taq PCR Master Mix (Ampliqon, Denmark) according to the manufacturer's instructions. PCR was performed under the following conditions: 94 °C for 3 min; followed by 35 cycles of 94°C for 30 s, 52°C for 30 s, and 72 °C for 90 s and ended with a final extension at 72 °C for 10 min. The expected PCR products (of ~700 base pairs) were purified and ligated into pTZ57R/T vector (Thermo Scientific, USA), according to the manufacturer's instructions. The ligation mix was transformed into *Escherichia coli* strain *DH5α*. Plasmid DNA from recombinant clones was purified using a Plasmid Miniprep Kit (Qiagen, Germany), and a purified clone from each isolate was subjected to sequencing in both directions (Macrogen Inc., South Korea). Sequence data were assembled using the Contig Express program in the Vector NTI 11 software (Invitrogen, USA).

Sequences, phylogenetic and recombination analysis

High nucleotide sequence similarity to ZYMV was indicated using BLAST N analysis. Analyses were conducted using 104 CI nucleotide sequences, including 10 nucleotide sequences obtained in this work and 94 retrieved from GenBank (Table 2). Out of ZYMV CI sequences retrieved from GenBank, three were from Iran and the others were from other countries in the world. The CI nucleotide sequences were translated to amino acids using ExPASy translate tool (http://web.expasy.org/translate/). Alignments were performed with Clustal W implemented in BioEdit v.7.2.5 (Hall, 1999). The pairwise nucleotide (nt) and amino acid (aa) sequence identity scores were displayed as color-coded cells using SDT v.1.2 software (Muhire et al., 2014). Phylogenetic trees were generated by the maximum-likelihood (ML) and neighbor-joining (NJ) methods implemented in MEGA7 (Kumar et al., 2016), with 1000 bootstrap replicates. Genetic distance between and within phylogenetic groups of ZYMV CI gene was calculated using MEGA7 with 1000 bootstrap replicates. Recombination analysis was performed on the aligned nucleotide sequences using RDP4 package (Martin et al., 2015). The occurrence of

Table 1. Characteristics of samples which identified as ZYMV after analyzing with BLAST and their origin, host, symptom and accession numbers

Isolates	Province (region)	Host plant	Symptom	Accession number
Gj1	Razavi Khorasan (Jovein)	*Solanum lycopersicum*	YM/MO	KJ135782
Kj	Mazandaran (Juybar)	*Cucurbita moschata*	M/B/Y	KJ135786
KHB2	Mazandaran (Babolsar)	*Cucurbita moschata*	M/MO/VB	KJ135785
KB	Mazandaran (Babolsar)	*Cucurbita moschata*	VC/M	KJ135784
KB2	Mazandaran (Babolsar)	*Cucurbita pepo*	VC/YSP	MF766014
Gj2	Razavi Khorasan (Jovein)	*Solanum lycopersicum*	MO/B/YM	MF766013
KS1	Mazandaran (Sari)	*Cucurbita pepo*	GB/D/M	MF766018
KF1	Mazandaran (Nowshahr)	*Cucurbita pepo*	M/B/D	MF766015
KG1	Golestan (Gorgan)	*Cucurbita pepo*	M/GB/D	MF766016
KG2	Golestan (Gorgan)	*Cucurbita pepo*	GB/D/S	MF766017

Abbreviations: VC; Vein clearing, VB; Vein banding, M; Mosaic, B; Blistering, GB; Green blistering, MO; Mottling; YSP; Yellow spots, YM; Yellow mosaic, Y; Yellowing, D; Deformation, S; Shoestring.

recombination events was assessed by at least four programs using default parameters, and a *P* value threshold of 0.05.

Population genetic analysis

Population genetic parameters of CI gene sequences obtained in this study and those from GenBank were estimated using DnaSP v. 6.10.04 software (Rozas et al., 2017) based on phylogenetic groups and geographic origins. Nucleotide sequences alignment of the CI gene were assessed to estimate number of haplotypes (H), haplotype diversity (Hd), number of polymorphic sites (S), total number of mutations (η), average pairwise nucleotide diversity (π) using the Jukes and Cantor correction (Jukes and Cantor, 1969), average number of nucleotide differences between sequences from the same population (K), and the ratio of non-synonymous to synonymous nucleotide diversity (dN/dS), also known as ω. In general, ω = 1, < 1 and > 1 indicates neutral evolution, negative (purifying) selection and positive (diversifying) selection, respectively. The nucleotide diversity measures the average pairwise variation among sequences with values ranging from 0 (no variation) to 0.1 (extreme variation). The haplotype diversity indicates the frequency of haplotypes in a sample with values ranging from 0 to 1.000 (Tsompana et al., 2005).

Population genetic differentiation

Genetic differentiation between populations was examined using several statistics: *Ks**, *Z*, *Z**, *Kst** and *Snn* based on permutation statistical tests with 1000 replicates. *Ks** and *Z* are the sequence-based statistics considered by Hudson (2000). Under the null hypothesis (no genetic differentiation), *Kst** is expected to be near zero, but if *Ks**, *Kst** test statistics is supported by small *P* value (<0.05), the null hypothesis is rejected (Hudson et al., 1992a). The *Z* statistic is calculated from ranking distances between all pairs of sequences. *Z** statistic is a logarithmic variant of *Z* statistic and if it is too small and supported by significant *P* value (<0.05) the null hypothesis of no genetic differentiation is rejected (Hudson et al., 1992b). The frequency of the nearest neighbor sequences in the same locality is measured by the *Snn* test statistic, whose values may range from 1 (when populations from different localities are genetically distinct) to 1/2 in the case of panmixia (Hudson, 2000). The degree of genetic differentiation or the level of gene flow between ZYMV populations was calculated by estimating the absolute value of the standardized variance in allele frequencies across populations (*Fst*) (Wright, 1951). The *Fst* values ranges from 0 (indicating no differentiation between the populations) to 1 (when the populations are clearly differentiated) (Rozas et al., 2003). These analyses were performed using DnaSP6 (Rozas et al., 2017).

Results

CI nucleotide sequences

PCR amplification of partial CI region yielded fragments of about 700 bp. Sequences of the CI gene from ten Iranian ZYMV isolates were successfully generated, submitted to the GenBank,

Table 2. GenBank accession number and origin of some of the previously reported ZYMV used for phylogenetic comparison of the nucleotides sequence of the CI coding region

Populations	Geography		Country	Number	Isolates/strains (host of origin)	Accession numbers
I	Asia		Iran	13	Fars (Cpe), IKA/strain A (Sq), SANRU (Cpe)	JN183062, KU528623, KU198853
			India	1	AP Gherkin (Can)	KT778297
			Turkey	35	YUN8-4 (Cpe), Y4 (Cpe), Y23 (Cpe), S3 (Cme), KZ1 (Cpe), KAR15-1 (Cmo), KAR12-4 (Cpe), K3 (Cpe), K17 (Cpe), H1M (Cs), G3 (Cpe), G2 (Cpe), G1 (Cpe), ER6-8 (Cpe), E-7 (Cme), AYS7 (Cpe), D14 (Cs), C5 (Cmo), C17 (Cme), C13 (Cmo), C11(Cmo), BRD4 (Cmo), BRD2 (Cmo), BE7 (Cmo), BE6 (Cpe), BE26 (Cpe), BE15 (Cpe), BE10 (Cpe), AS5 (Cpe), AS1 (Cpe), AS11 (Cpe), AKS6-2 (Cpe), AKS5-7 (Cpe), AKS2-5 (Cmo), KZN1 (Cpe)	KP828427, KP828426, KP828425, KP828424, KP828423, KP828422, KP828421, KP828420, KP828419, KP828418, KP828417, KP828416, KP828415, KP828414, KP828413, KP828412, KP828411, KP828410, KP828409, KP828408, KP828407, KP828406, KP828405, KP828403, KP828402, KP828400, KP828397, KP828395, KP828394, KP828393, KP828392, KP828391, KP828390, KP828389, KP828388
			Israel	3	AG (-), NAT (-), B* (France-Israel) (-)	EF062583, EF062582, AY188994
			South Korea	5	RDA (Cpe), KR-PS (Cmo),KR-PE (Cmo), KR-PA (Cmo), A (Ar)	AB369279, AY279000, AY278999, AY278998, AJ429071
			Japan	4	2002 (Cs), Z5-1 (Cs), 169 (Cme), M (-)	AB188116, AB188115, AB020477, AB020478
			China	9	WS (Cpe), zz (Sin), SXSG (La), CJLX30535 (Crayfish), spider131932 (Spiders), WG (Bh), SG (Lc), CU (Cs), WM (Cl)	KX664482, KX421104, KX249747, KX884565, KX884570, AJ316229, AJ316228, AJ307036, AJ515911
			Taiwan	3	TW-TN3 (Lc), Begonia (Begonia), TW-TN3 (Lc)	NC_003224, AM422386, AF127929
	Africa		Egypt	1	EG (Sq)	LC153708
	Europe		Slovakia	2	Kuchyna (Cpe), SE04T (Cpe),	DQ124239, KF976713
			Czech Republic	1	H (Cpe)	KF976712
			Spain	1	Vera (Cpe)	KX499498
	Americas	N. America	USA	11	leaf23 (Cpe), leaf17 (Cpe), leaf1 (Cpe), SG5 (Cpe), SG4 (Cpe), SG1 (Cpe), FG2 (Cpe), PA_2006 (Cpe), California (Cmo), - (Cpe), - (Cpe)	KJ923769, KJ923768, KJ923767, KC665635, KC665634, KC665631, KC665630, JQ716413, L31350, KJ875864, KJ875865

		S. America	Argentina	1	10itSDE (Cma)	KT598222
	Oceania		Australia: Broome, WA	3	13Br (Cpe), 20Br (Cpe), 56Br (Cpe)	KY225555, KY225550, KY225549
			Australia: Darwin, NT	2	38NT (Cme, honeydew), 75NT (Cme, rockmelon)	KY225548, KY225547
II	Asia		Singapore	2	Singapore (-), Singapore (Cs)	U60962, AF014811
	Africa		Reunion Island	1	Reunion Island (Mch)	L29569
	Oceania		East Timor	3	TM40 (Cs), TM16 (Pu), TM39 (Pu)	KY225556, KY225545, KY225544
III			Australia: Kununurra, WA	3	694K (Pu), 695K (P), 697K (Cme, honeydew)	KY225543, KY225542, KY225546

Cpe: *Cucurbita pepo*, Cme: *Cucumis melo*, Cma: *Cucurbita maxima*, Cmo: *Cucurbita moschata*, Cs: *Cucumis sativus*, Can: *Cucumis anguria*, Mch: *Momordica charantia,* Cl: *Citrullus lanatus*, Sin: *Sesamum indicum,* La: *Luffa aegyptiaca*, Lc: *Luffa cylindrica, Bh: Benincasa hispida,* Ar*: Althaea rosea,* Pu: Pumpkin, Sq: Squash, -: unknown isolate or host

and assigned the accession numbers KJ135782, KJ135784, and MF766013-MF766018 (Table 1). Names and accession numbers of the previously reported ZYMV isolates have been also presented in Table 2.

Sequence comparisons

The pairwise sequence identity of partial CI gene of all 104 ZYMV isolates ranged from 79.0 to 100.0% at the nt sequence level (Figure 1) and from 91.2 to 100% at the aa sequence level (Figure S1). All 13 Iranian isolates (ten from this study and three retrieved from GenBank) revealed 93.5–99.1% and 94.3–100% identity at the nt and aa levels, respectively. The lowest nt identity (79.0%) was observed between Gj1 and TM40 (KY225556) and TM39 (KY225544) isolates from East Timor. In addition, the highest nt identity (99.7%) was identified between KG1 and AG, NAT and B isolates from Israel. Amino acid sequence identity in the CI gene of all ZYMV isolates was over 91%. The minimum aa sequence identity of the CI gene between the Iranian isolates and those deposited in GenBank was between isolates Gj1, Gj2 and KB and isolate TM39 (KY225544, East Timor) (91.2%), respectively. Some Iranian isolates (Fars, SANRU, KS1, KG1, KG2, KF1) showed 100% aa identity with isolates from Slovakia (SE04T, Kuchyna), Czech Republic (H), Israel (NAT, B, AG), Japan (169) and Turkey (KZN1, AKS2-5, AKS5-7, AKS6-2, AS11, AS5, BE10, BE15, BE26, BE6, BE7, BRD4, C11, C13, C17, C5, D14, E-7, ER6-8, G1, G2, G3, H1M, K3, KZ1, KAR12-4, KAR15-1, S3).

Phylogenetic analysis

The ZYMV CI coding region sequences were subjected to phylogenetic analyses, with that of *Watermelon mosaic virus* (WMV) isolate (IR02-54, EU660584) as outgroup. Both the ML and NJ trees showed a similar topology. As shown in Figure 1, all the 104 ZYMV isolates were divided into three distinct phylogroups: I, II and III. Group I is a large and geographically widespread group which was further clustered into several subgroups (IA, IB, IC and ID). Group I included a range of isolates (n=95) from different parts of the world including all 13 Iranian isolates plus isolates from Egypt (n=1), Turkey (n=35), Australia (n=5), Argentina (n=1), USA (n=11), Spain (n=1), India (n=1), Slovakia (n=2), Czech Republic (n=1), Israel (n=3), Taiwan (n=3), Japan (n=4), South Korea (n=5), and China (n=9). The between-subgroup genetic distance of the four subgroups in group I was significantly higher than the within-subgroup ones (Table S1) which providing evidence for a phylogenetic grouping. The overall mean value of nucleotide sequence diversity between Iranian and other isolates in subgroup IA was 0.031±0.003. Group II included two isolates from Singapore. Group III contained three isolates from East Timor, one from Reunion Island and three from Australia. The overall mean distance among all ZYMV isolates was 0.077±0.006. Based on pairwise comparisons, genetic distance within groups was 0.052±0.004, 0.000±0.000 and 0.130±0.010 for

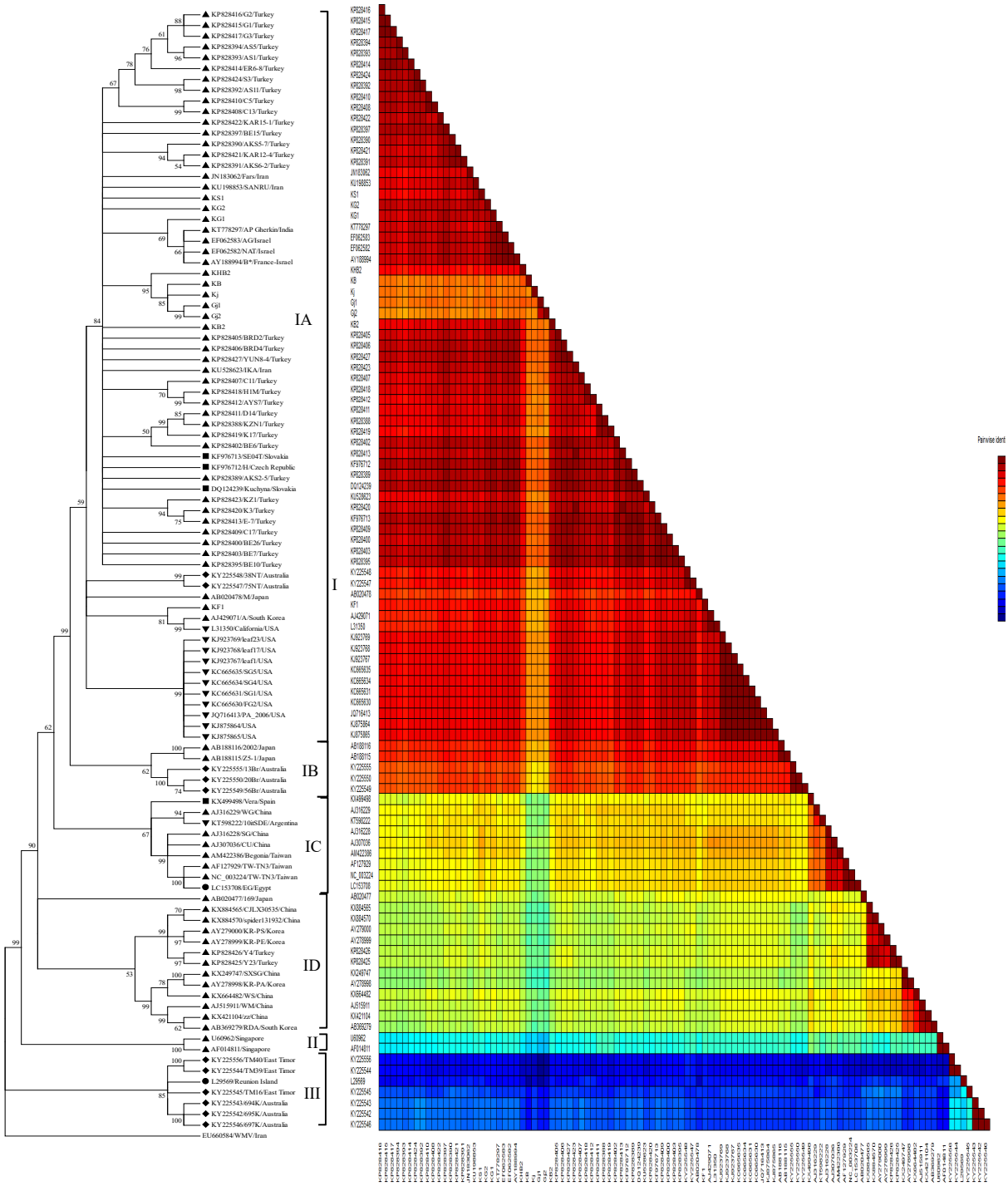

Figure 1. Maximum-likelihood phylogenetic tree constructed from the partial cylindrical inclusion (CI) gene nucleotide sequences of 104 ZYMV isolates, and graphical representation of pairwise nucleotide identity. The phylogenetic tree was generated in MEGA7 and bootstrapped with 1000 replicates. Bootstrap values ≥ 50% are shown at the branch internodes. Two dimensional nucleotide diversity plot constructed based on SDT MUSCLE alignment. The Asian isolates are indicated by "▲", European isolates by "■", American isolates by "▼", Oceanian isolates by "♦" and African isolates by "●".

group I, II and III, respectively. In addition, the genetic diversity between groups was 0.144±0.015, 0.220±0.018 and 0.231±0.019 for group I versus II, I versus III and II versus III, respectively. As expected, the genetic distances between the three groups were significantly greater than the within-group ones, supporting the results of phylogenetic grouping.

Table 3. Population genetic parameters calculated for the CI genes of *Zucchini yellow mosaic virus* on the basis of phylogroups identified in Figure 1 and geographical origin

Phylogroup	H	Hd	S	η	K	π	SS	NS	dN	dS	ω
All	72	0.986	280	373	46.46	0.074	154.67	526.33	0.00784	0.27345	0.0286
Group I (n=95)	67	0.985	**234**	**278**	33.026	0.050	154.64	526.36	0.00549	0.2496	**0.0219**
Group II (n=2)	1	0.000	0	0	0	0.000	153.67	527.33	0.0000	0.0000	0.0000
Group III (n=7)	4	0.810	172	179	**77.619**	**0.126**	155.38	525.62	0.01117	0.87953	0.0127
Geographic origins											
Asia (n=75)	60	0.994	**236**	**288**	37.767	0.058	154.68	526.32	0.00648	0.30088	**0.02153**
Europe (n=4)	2	0.500	53	53	26.500	0.041	154.50	526.50	0.00382	0.20613	0.01853
America (n=12)	3	0.318	55	56	10.000	0.015	154.47	526.53	0.00063	0.07950	0.0079
Oceania (n=11)	7	0.909	194	211	89.127	0.148	154.67	526.33	0.01605	1.70054	0.00943
Africa (n=2)	2	1.000	126	126	**126.000**	**0.212**	156.17	524.83	0.01930	3.48410	0.005539

H, number of haplotypes, Hd, haplotype diversity; S: number of polymorphic sites; η (eta): total number of mutations; k: average number of nucleotide differences between sequences; π: nucleotide diversity, with Jukes & Cantor correction; SS: total number of synonymous sites analyzed; NS: total number of non-synonymous sites analyzed; dN, average number of nonsynonymous substitutions per nonsynonymous site; dS, average number of synonymous substitutions per synonymous site, with the Jukes and Cantor correction; dN/dS, average ratio between nonsynonymous and synonymous substitutions in sequence pairs. Maximum respective values between groups are in bold.

It is worth noting that no recombination event was found between ZYMV isolates in CI gene. Also, no signatures of recombination were detected between ZYMV group I, II and III subpopulations, indicating significant genetic differentiation and limited gene flow between isolates in these phylogenetic groups, probably due to the presence of quarantine and physical barriers between them or existence of other host plants.

Genetic diversity of ZYMV

Pairwise comparisons showed that members of group I shared 87.7-100% nt sequence identity, with an average nt identity value of 97.05%, members of group II were 100% identical, and members of group III shared 84.6-100% nt sequence identities, with an average nt identity value of 92.68% (Figure 1). This suggested that group III ZYMVs had a higher level of genetic variation than those belonging to group I and II. Group I did not show a clear division in terms of geographical distribution. However, groups II and III were more phylogenetically clustered by geographical origin. Haplotype diversity and nucleotide diversity for all ZYMV isolates were 0.986 and 0.074, respectively, indicating a relatively high genetic diversity in ZYMV populations and among lineage subpopulations (Table 3). The haplotype diversity for group I and group III was 0.985 and 0.810, whereas nucleotide diversity for these two groups was 0.050 and 0.126, respectively. Notably, it was impossible to perform these statistical tests for ZYMV group II isolates, due to limited data. The highest nucleotide diversity (π=0.126) between the isolates and the greatest overall average number of differences, k (78 nucleotides), were calculated for the phylogroup III. However, the largest number of segregation sites, (S=234), and mutations within the segregating sites, (η=278), were found in the phylogroup I (Table 3). In the geographical populations, the highest values of π (=0.212) and k (126 nucleotides), were calculated for the Africa population. However, the highest values of S (236), and η (288), were found in Asia population (Table 3). The lowest π (0.015) and k (10 nucleotides) were estimated for the America population. The global selection pressure (dN/dS) for all ZYMV isolates was 0.0286. Furthermore, the dN to dS ratio (ω) for each population was <1. The highest

and lowest pressure was calculated for Asian (ω=0.021) and African (ω=0.005) populations. These results indicated that all ZYMV populations are under negative selection but subjected to distinct constraints. To determine the gene- and site-specific selection pressures acting on the ZYMV CI cistron, different codon-based maximum-likelihood algorithms within the HYPHY software package as implemented in Datamonkey server (www.datamonkey.org) were used to estimate the value of ω at each codon site. All of the codons were under negative selection or neutral evolution, which revealed that strong purifying evolutionary constraint is driving CI gene evolution in ZYMV.

Differentiation of phylogroups and geographical populations

As mentioned, genetic distinction of ZYMV populations was defined in two categories: phylogenetic populations and geographical populations. With exception of insignificant *Snn* value for group II vs. III, the independent statistical tests of population differentiation (*Ks**, *Kst**, *Z** and *Snn*) were significant (Table 4), supporting the genetic differentiation between lineage groups of ZYMV isolates. Strong genetic differentiation confirmed by high F_{ST} (>0.549). Additionally, gene flow and genetic differentiation between the Asian, American, European, Oceanian and African populations of the ZYMV isolates were determined using the *Ks**, *Kst**, *Z**, *Snn* and F_{ST} statistical tests. Among the ZYMV geographical populations, American and Oceanian populations with significant *Kst**, high Snn (mostly near 1.000) and F_{ST} (0.352) values are statistically distinct. However, nonsignificant *Ks**, *Kst**, *Z** and *Snn* values were indicated no significant differentiation between European population with the Asian and the African populations. Such a nonsignificant differentiation was also associated with low F_{ST} value (<0.104). Genetic differentiation between Asia vs. America, Asia vs. Oceania, and Europe vs. America confirmed by *Kst**, *Z**, *Snn*, and relatively high F_{ST} value (0.223-0.232), suggesting significant genetic differentiation. Also, no genetic differentiation was observed between Asia vs. Europe and Oceania vs. Africa, due to negative F_{ST} values or nonsignificant *Ks**, *Kst**, *Z** and *Snn* values (Table 4). Genetic isolation was less pronounced between the African population with the Asian and the European populations, indicating frequent gene flow. There was frequent gene flow between the European ZYMV with Oceanian and American with African populations, because the related F_{ST} values were <0.33. In addition, nonsignificant *Z*, *Z** or *Snn* values indicated these population pairs were not well differentiated (Table 4). Taken together, there is some significant correlation between geographical position and genetic distances among the geographical populations, showing that observed genetic differentiation could be explained by distance isolation.

Discussion

Analysis of genetic variation in ZYMV populations from different geographical locations can provide relevant information for understanding its emergence, epidemiology, and gene flow. Phylogenetic analysis and genetic differentiation of 104 ZYMV isolates, revealed that the population structure of the three ZYMV phylogroups somewhat correlated with their geographical locations; which was supported by the subsequent genetic distance analyses. Previous ZYMV studies used complete or partial CP sequences to distinguish phylogenetic groups. Desbiez et al. (2002) classified ZYMV isolates into two main groups based on the analyses of 47 partial nt sequences of CP gene. After analyzing the complete CP nt sequences of 39 ZYMV isolates, Zhao et al. (2003) designated three groups (I-III): I, worldwide; II, containing isolates only from Asia; and III, containing isolates only from China. Subsequently, Ha et al. (2008a) analyzed the complete CP nt sequences of 61 ZYMV isolates into three main clusters: I, distributed worldwide; II, comprising Reunion Island, Singapore and Vietnam isolates; and III, consisting of Vietnam and China isolates. By comparison of 208 partial CP sequences (231 nt), Bananej et al. (2008) suggested two main groups. Group A was a worldwide group that included three subgroups, and B comprised isolates from China, Reunion Island, Singapore and Vietnam. By analyzing the 143 complete CP sequences, Coutts et al. (2011) classified ZYMV isolates into three main groups as proposed by Ha et al. (2008a). Similarly, Massumi et al. (2011) got the same results in analyses based on the nucleotide sequences of the whole CP gene and the NIb-CP gene fragment. Finally, Maina et al. (2017) analyzed ZYMV populations from East Timorese and northern Australia and found connectivity

Table 4. Results of genetic differentiation analysis between subpopulations from pairwise comparison of *Zucchini yellow mosaic virus* sequences based on phylogroups identified in Figure 1 and geographical populations

Comparisons	*Ks** (*P*-value)	*Kst** (*P*-value)	*Z* (*P*-value)	*Z** (*P*-value)	*Snn* (*P*-value)	F_{ST}
Phylogroup						
Group I vs. II	3.217 (0.000)	0.015 (0.000)	2233.150 (0.000)	7.411 (0.000)	1.000 (0.003)	0.809
Group I vs. III	3.242 (0.000)	0.054 (0.000)	2307.149 (0.000)	7.418 (0.000)	1.000 (0.000)	0.549
Group II vs. III	3.700 (0.0450)	0.087 (0.045)	10.904 (0.033)	2.293 (0.026)	1.000 (0.056 ns)	0.700
Geography						
Asia vs. Europe	3.295 (0.559 ns)	-0.001 (0.559 ns)	1559.699 (1.000 ns)	7.041 (0.525 ns)	0.875 (0.875 ns)	-0.041
Asia vs. America	3.054 (0.000)	0.067 (0.000)	1765.184 (0.001)	6.938 (0.000)	0.942 (0.000)	0.223
Asia vs. Oceania	3.398 (0.000)	0.039 (0.000)	1698.290 (0.000)	7.061 (0.000)	0.994 (0.000)	0.232
Asia vs. Africa	3.331 (0.015)	0.016 (0.015)	1414.834 (0.020)	6.949 (0.017)	0.961 (0.253 ns)	0.085
Europe vs. America	1.192 (0.000)	0.355 (0.000)	47.472 (0.002)	3.656 (0.00)	0.875 (0.002)	0.226
Europe vs. Oceania	3.582 (0.035)	0.078 (0.035)	50.564 (0.292 ns)	3.567 (0.066 ns)	1.000 (0.001)	0.263
Europe vs. Africa	1.994 (0.138 ns)	0.419 (0.138 ns)	4.500 (0.138 ns)	1.445 (0.138 ns)	0.500 (0.584 ns)	0.104
America vs. Oceania	2.407 (0.000)	0.246 (0.000)	102.046 (0.000)	4.275 (0.000)	0.956 (0.000)	0.352
America vs. Africa	1.032 (0.025)	0.470 (0.014)	34.500 (0.025)	3.447 (0.019)	0.786 (0.138 ns)	0.208
Oceania vs. Africa	3.935 (0.019)	0.047 (0.019)	36.809 (0.093 ns)	3.308 (0.061ns)	0.846 (0.097 ns)	-0.058

Note: Probability (*P*-value) obtained by the permutation test (PM test) with 1000 replicates. ns, not significant. The analysis was done using DnaSP v. 6.10.04.

between them either in the genome-based tree or the CP-based tree. In this study, pairwise comparisons and phylogenetic analysis based on partial CI gene nt sequences clearly showed the existence of three groups, in which phylogroups I (worldwide) and III (East Timor, Reunion Island and Australia-Kununurra) were consistent with the genomic nt sequence phylogroup classification of Maina et al. (2017) but phylogroup II (containing only two Singapore isolates) was an additional one. Group I was the largest and widespread group including most of the ZYMV isolates from Asia, Europe, North and South America, Africa and Australia, in accordance with previous reports by Desbiez et al. (2002), Bananej et al. (2008), Coutts et al. (2011) (they denoted group I as group A), Ha et al. (2008a), Massumi et al. (2011) and Maina et al. (2017). This study also suggested four minor groups within group I, in which subgroups A, B, C, and D corresponded to the reported subgroups II, I, IV+V, and III, respectively (Maina et al., 2017). However, subgroup IV along with the previously reported WG and 10itSDE isolates in subgroup V were integrated into subgroup IC. As mentioned above, the geographical origins of the isolates in group I were the most diverse and the overall nt and aa identity within CI sequences in this group was >87.0% and >92.0%, respectively, which suggest the common origin of distantly distributed isolates. International trading of infected seeds, plants or fruits can be a possible explanation for such sequence similarities observed between the intercontinental isolates of ZYMV (Desbiez et al., 2002; Lecoq et al., 2003; Simmons et al., 2008; Simmons et al., 2011, 2013). In some cases, the CI gene data was more phylogenetically classified by geographical situation than anticipated by chance alone, as depicted in subgroups IC (expect Argentinian isolate) and ID. Analysis of ZYMV population differentiation indicated that three phylogroups were completely distinct with significant *Ks**, *Kst**, *Z*, Snn* and very high F_{ST} values (>0.500). The ω estimates for group I and group III were respectively 0.022 and 0.013 (Figure 1, Table 3), and in concordance with the result of genetic differentiation analysis (Table 4). The result showed that group I was subjected to more intense purifying selection than group III. Recombination is one of the principal forces driving plant virus evolution (Garcia-Arenal et al., 2003), however no recombination event was detected in CI gene of studied isolates, suggesting that this potent evolutionary force has not shaped the emergence of ZYMV CI gene variants. Meanwhile, the partial genome fragment could not provide accurate results. A previous study provided evidence for the presence of recombination cold spots within the full-length polyprotein of 14 ZYMV isolates from northern Australia (n=10, Broome, Kununurra), East Asia (n=2, Japan, China) and Southeast Asia (n=2, Singapore, East Timor) (Maina et al., 2017). Among them, Z5-1 from Japan was lone isolate identified as a recombinant in the CI coding region plus 6k2, NIa-Vpg and NIa-Pro coding regions and the lower frequency of recombination occurred in these regions than elsewhere in genomic RNA. Overall, there were a low frequency of recombination in most of ZYMV isolates (Maina et al., 2017); one possible explanation is strong selective pressure against survival of new ZYMV recombinants. In genetic diversity analyses (Table 3), the African population showed the most nucleotide diversity (π), followed by the Oceanian population. However, American and European populations exhibited low haplotype diversity (0.318, 0.500) and nucleotide diversity (0.015, 0.041). Low level of genetic diversity among American isolates as well as European ZYMV isolates was in contrast to the diversity reported from other parts of the world. Geographical population cluster levels of genetic differentiation ranged from -0.041–0.352 in the F_{ST} values. The highest and lowest F_{ST} values were found for Oceania versus America and Africa populations, respectively (Table 4). Except African population, all the populations were differentiated from the American and Oceanian ZYMV populations because the *Kst** values were well above zero and supported by high significant *P*-value (0.000). The extent of genetic differentiation between most of the geographical population pairs was moderate ($0.085<F_{ST}<0.104$) to great ($0.223<F_{ST}<0.263$), indicating moderate to high gene flow between these geographical

ZYMV populations. The exception was Oceanian and American ZYMV populations, which had complete genetic difference (infrequent gene flow) (F_{ST} =0.352). Genetic differentiation between American and Oceanian ZYMV populations also confirmed by all statistical tests. This could be due to long distances between these geographic regions, indicative of a correlation between genetic and geographical distances. Based on these test statistics, geographical isolation may have played a role in ZYMV population structure especially in Oceanian and American isolates. The dN/dS ratio for Asian isolates was the highest, indicating that CI is under tighter functional constraints for these isolates. There were no codons identified as being under positive selection for all lineages. Strong negative selection on the CI of the ZYMV suggests the crucial role of this protein in helicase activities, RNA replication, cell-to-cell and systemic movement or other vital yet unknown functions (Carrington et al., 1998; Klein et al., 1994). In the phylogenetic analysis all ZYMV populations were polyphyletic and distributed in more than one phylogenetic groups (Figure 1). This indicates that ZYMV isolates were dispersed to other geographical areas with unknowingly infected seed (despite low levels of seed transmission) (Tobias and Palkovics, 2003; Desbiez and Lecoq, 1997; Schrijnwerkers et al., 1991; Simmons et al., 2011) or vegetative propagules and evolved via genetic drift (founder effect). As mentioned, the sequence variation along the CI gene of ZYMV isolates is controlled by purifying selection pressure (<1). Alternatively, in situ evolution within several countries, with human activity in widespread seed transmission playing a main role in ZYMV dispersal, as suggested by Simmons et al. (2008) in analysis of ZYMV CP gene. Therefore, when an isolate becomes settled down in a place, without positive selection within the population, little change could occur unless a new variant is introduced, as the case for Australian isolates (Kununurra in northern Australia which are highly different from other Australian isolates) (Coutts et al., 2011). Moreover, the Kununurra sequences grouped together with the three East Timorese sequences within major phylogroup III (previously called the Southeast Asian/Reunion Island phylogroup), which seems adapted to tropical conditions (Maina et al., 2017). The close relationship between the CI sequences (as well as complete genomic sequences) from Kununurra and East Timor suggest recent ZYMV introduction across the sea from Southeast Asia to Kununurra. Such grouping could be attributed to monsoonal winds (from East Timor toward northern Australia) which could bring viruliferous insect vectors or migrating birds with infected seed in their guts, thus introducing viruses (Eagles et al., 2013). The Iranian ZYMV isolates in the subgroup IA shared 93.5–99.1% CI nucleotide sequence identity with each other and 87.7-99.7% with other isolates of this subgroup. Iranian isolates were more resembling to isolates from Middle East (Israel, Turkey and India,), Far East (China, Japan and South Korea), Europe (Spain, Czech Republic, and Slovakia), Australia and USA in partial CI nucleotide sequence. So, how ZYMV first entered Iran is difficult to determine, but there are several possible pathways. During commercial exchanges, infected cucurbit material such as plants, fruits or seeds may have entered from elsewhere, providing the initial virus source. In the present study, tomato was found to be a new natural host of ZYMV, broadening the understanding of the genetic diversity of the pathogen in pathogenicity to plants. The analyses done in this study provide evidence for important evolutionary forces driving ZYMV evolution such as selection, genetic drift and founder effects by exchange infected plant products between different geographical regions. These findings provide an insight into the ZYMV population structure and are helpful for designing proper strategies to the management of this virus.

References

1. Adams M.J., Antoniw J.F. and Beaudoin F. (2005a) Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. Molecular Plant Pathology 6:471–487.
2. Adams M.J., Antoniw J.F. and Fauquet C.M. (2005b) Molecular criteria for genus

and species discrimination within the family *Potyviridae*. Archives of Virology 150:459–479.
3. Adams M.J., Zerbini F.M., French R., Rabenstein F., Stenger D.C. and Valkonen J.P.T. (2012) Family *Potyviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, eds. Virus taxonomy: Classification and nomenclature of viruses. Ninth report of the ICTV. London, UK: Academic Press, 1069–1089.
4. Al-Musa A.M. (1989) Over summering hosts for some cucurbit viruses in the Jordan Valley. Journal of Phytopathology 127:49–54.
5. Bananej K.A., Keshavarz T., Vahdat A., Hosseini Salekdeh G. and Glasa M.J. (2008) Biological and molecular variability of *Zucchini yellow mosaic virus* in Iran. Journal of Phytopathology 156:654-659.
6. Chen Y.K. and Hong Y.H. 2008. First report of Begonia chlorotic ringspot caused by *Zucchini yellow mosaic virus* in Taiwan. Plant Disease 92:1247.
7. Choi S.K., Yoon J.Y., Ryu K.H., Choi J.K. and Park W.M. (2002) First report of *Zucchini yellow mosaic virus* on hollyhock (*Althaea rosea*). Plant Pathology Journal 18:121–125.
8. Coutts B.A. and Jones R.A.C. (2005) Incidence and distribution of viruses infecting cucurbit crops in the Northern Territory and Western Australia. Australian Journal of Agricultural Research 56:847–858.
9. Coutts B.A., Kehoe M.A., Webster C.G., Wylie S.J. and Jones R.A. (2011) *Zucchini yellow mosaic virus*: biological properties, detection procedures and comparison of coat protein gene sequences. Archives of Virology 156:2119-31.
10. Chung BY-W., Miller W.A., Atkins J.F. and Firth A.E. (2008) An overlapping essential gene in the *Potyviridae*. Proceedings of the National Academy of Sciences of the USA 105:5897–5902.
11. Desbiez C., Wipf-Scheibel C., Granier F., Robaglia C., Delaunay T. and Lecoq H. (1996) Biological and molecular variability of *Zucchini yellow mosaic virus* in the island of Martinique. Plant Disease 80:203–207.
12. Desbiez C. and Lecoq H. (1997) *Zucchini yellow mosaic virus*. Plant Pathology 46: 809-829.
13. Desbiez C., Wipf-Scheibel C. and Lecoq H. (2002) Biological and serological variability, evolution and molecular epidemiology of *Zucchini yellow mosaic virus* (ZYMV, *Potyvirus*) with special reference to Caribbean islands. Virus Research 85:5–16.
14. Eagles D., Walker P.J., Zalucki M.P. and Durr P.A. (2013) Modelling spatio-temporal patterns of long-distance Culicoides dispersal into northern Australia. Preventive Veterinary Medicine 110: 312-322.
15. Gal-On A. (2007) *Zucchini yellow mosaic virus*: insect transmission and pathogenicity-the tails of two proteins. Molecular Plant Pathology 8:139–150.
16. Garcı´a-Arenal F., Fraile A. and Malpica J.M. (2003) Variation and evolution of plant virus populations. International Microbiology 6:225–232.
17. Gibbs A. and Ohshima K. (2010) Potyviruses and the digital revolution. Annual Review of Phytopathology 48: 205–223.
18. Glasa M., Svoboda J. and Novakova S. (2007) Analysis of the molecular and biological variability of *Zucchini yellow mosaic virus* isolates from Slovakia and Czech Republic. Virus Genes 35:415–421.
19. Ha C., Revill P., Harding R.M., Vu M. and Dale J.L. (2008a) Identification and sequence analysis of potyviruses infecting crops in Vietnam. Archives of Virology 153:45–60.
20. Ha C., Coombs S., Revill P.A., Harding R.M., Vu M. and Dale J.L. (2008b) Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Archives of Virology 153:25-36.
21. Hall T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98.
22. Hudson R.R., Boos D.D. and Kaplan N.L. (1992a) A statistical test for detecting geographic subdivision. Molecular Biology and Evolution 9:138–151.
23. Hudson R.R., Slatkin M. and Maddison W.P. (1992b) Estimation of levels of gene flow from DNA sequence data. Genetics 132:583–589.
24. Hudson R.R. (2000) A new statistic for detecting genetic differentiation. Genetics 155: 2011–2014.

25. Carrington J.C., Jensen P.E. and Schaad M.C. (1998) Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. Plant Journal 14: 393–400.
26. Jukes T.H. and Cantor C.R. (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, NewYork, pp 21–132.
27. Klein P.G., Klein R.R., Rodriguez-Cerezo E., Hunt A.G. and Shaw J.G. (1994) Mutational analysis of the *Tobacco vein mottling virus* genome. Virology 204:759-769.
28. Kumar S., Stecher G., and Tamura K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.
29. Lecoq H., Pitrat M. and Clément M. (1981) Identification et caractérisation d'un potyvirus provoquant la maladie du rabougrissement jaune du melon. Agronomie 1:827–834.
30. Lecoq H., Bourdin D., Raccah B., Hiebert E. and Purcifull D.E. (1991) Characterization of a *Zucchini yellow mosaic virus* isolates with a deficient helper component. Phytopathology 81: 1087-1091.
31. Lecoq H., Desbiez C., Wipf-Scheibel C. and Girard M. (2003) Potential involvement of melon fruit in the long-distance dissemination of cucurbit potyviruses. Plant Disease 87:955–959.
32. Lee K.C., Mahtani P.H., Chng C.G. and Wong S.M. (1997) Sequence and phylogenetic analysis of the cytoplasmic inclusion protein gene of *Zucchini yellow mosaic potyvirus*: its role in classification of the *Potyviridae*. Virus Genes 14: 41–53.
33. Lisa V., Boccardo G., D'Agostino G., Dellavalle G. and d'Aquilio M. (1981) Characterization of a potyvirus that causes *Zucchini yellow mosaic virus*. Phytopathology 71: 667-672.
34. Lisa V. and Lecoq H. (1984) *Zucchini yellow mosaic virus*. CMI/AAB Description of Plant Viruses, No. 282.
35. Martin D.P., Murrell B., Golden M., Khoosal A. and Muhire B. (2015) RDP4: detection and analysis of recombination patterns in virus genomes. Viral evolution 1:vev003.
36. Massumi H., Shaabanian M., Heydarnejad J., Hosseini Pour A. and Rahimian H. (2011) Host range and phylogenetic analysis of Iranian isolates of *Zucchini yellow mosaic virus*. Journal of Plant Pathology 93:187-193.
37. Maina S., Coutts B.A., Edwards O.R., Almeida L., Kehoe M.A., Ximenes A., and Jones R.A.C. (2017) *Zucchini yellow mosaic virus* populations from East Timorese and Northern Australian Cucurbit Crops: Molecular Properties, Genetic Connectivity, and Biosecurity Implications. Plant Disease 101:1236-1245.
38. Moury B., Morel C., Johansen E. and Jacquemond M. (2002) Evidence for diversifying selection in *Potato virus Y* and in the coat protein of other potyviruses. Journal of General Virology 83:2563-2573.
39. Muhire B.M., Varsani A. and Martin D.P. (2014) SDT: A Virus Classification Tool Based on Pairwise Sequence Alignment and Identity Calculation. PLoS ONE 9: e108277.
40. Novakova S., Svoboda J. and Glasa M. (2014) Analysis of the complete sequences of two biologically distinct *Zucchini yellow mosaic virus* isolates further evidences the involvement of a single amino acid in the virus pathogenicity. Acta Virologica 58:364 – 367.
41. Riechmann J.L., Lain S. and Garcia J.A. (1992) Highlights and prospects of *Potyvirus* molecular biology. Journal of General Virology 73:1–16.
42. Rozas J., Sanchez-DeI Barrio J.C., Messeguer X. and Rozas R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497.
43. Rozas J., Ferrer-Mata A., S anchez-DelBarrio J.C., Guirao-Rico S., Librado P., Ramos-Onsins S.E. and Sanchez-Gracia A. (2017) DnaSP 6: DNA sequence polymorphism analysis of large datasets. Molecular Biology and Evolution 34:3299-3302.
44. Schrijnwerkers C.C.F.M., Huijberts N. and Bos L. (1991) *Zucchini yellow mosaic virus*; two outbreaks in the Netherlands and seed transmissibility. Netherlands Journal of Plant Pathology 97:187–191.
45. Simmons H.E., Holmes E.C. and Stephenson A.G. (2008) Rapid evolutionary dynamics of *Zucchini yellow mosaic virus*. Journal of General Virology 89:1081–1085.

46. Simmons H.E., Holmes E.C., Gildow F.E., Bothe-Goralczyk M.A. and Stephenson A.G. (2011) Experimental Verification of Seed Transmission of *Zucchini yellow mosaic virus*. Plant Disease 95:751–754.
47. Simmons H.E., Dunham J.P., Zinn K.E., Munkvold G.P., Holmes E.C. and Stephenson A.G. (2013) *Zucchini yellow mosaic virus* (ZYMV, *Potyvirus*): Vertical transmission, seed infection and cryptic infections. Virus Research 176:259–264.
48. Tobias I. and Palkovics L. (2003) Characterization of Hungarian isolates of *Zucchini yellow mosaic virus* (ZYMV, *Potyvirus*) transmitted by seeds of *Cucurbita pepo* var. Styriaca. Pest Management Science 59:493–497.
49. Tsompana M., Abad J., Purugganan M. and Moyer J.W. (2005) The molecular population genetics of the *Tomato spotted wilt virus* (TSWV) genome. Molecular Ecology 14:53–66.
50. Urcuqui-Inchima S., Haenni A. and Bernardi F. (2001) *Potyvirus* proteins: a wealth of functions. Virus Research 74:157-175.
51. Wright S. (1951) The genetical structure of populations. Annals of eugenics 15:323–354.
52. Yakoubi S., Desbiez C., Fakhfakh H., Wipf-Scheibel C., Fabre F., Pitrat M., Marrakchi M. and Lecoq H. (2008) Molecular, biological and serological variability of *Zucchini yellow mosaic virus* in Tunisia. Plant Pathology 57:1146–1154.
53. Zhao M.F., Chen J., Zheng H.Y., Adams M.J. and Chen J.P. (2003) Molecular analysis of *Zucchini yellow mosaic virus* isolates from Hangzhou, China. Journal of Phytopathology 151:307-311.

Table S1. Genetic distances within and between subgroups in group I

	Subgroups			
	IA	**IB**	**IC**	**ID**
Subgroup IA	0.022 ± 0.002			
Subgroup IB	0.038 ± 0.005	0.019 ± 0.004		
Subgroup IC	0.079 ± 0.009	0.080 ± 0.010	0.034 ± 0.004	
Subgroup ID	0.099 ± 0.011	0.102 ± 0.011	0.090 ± 0.009	0.057 ± 0.006

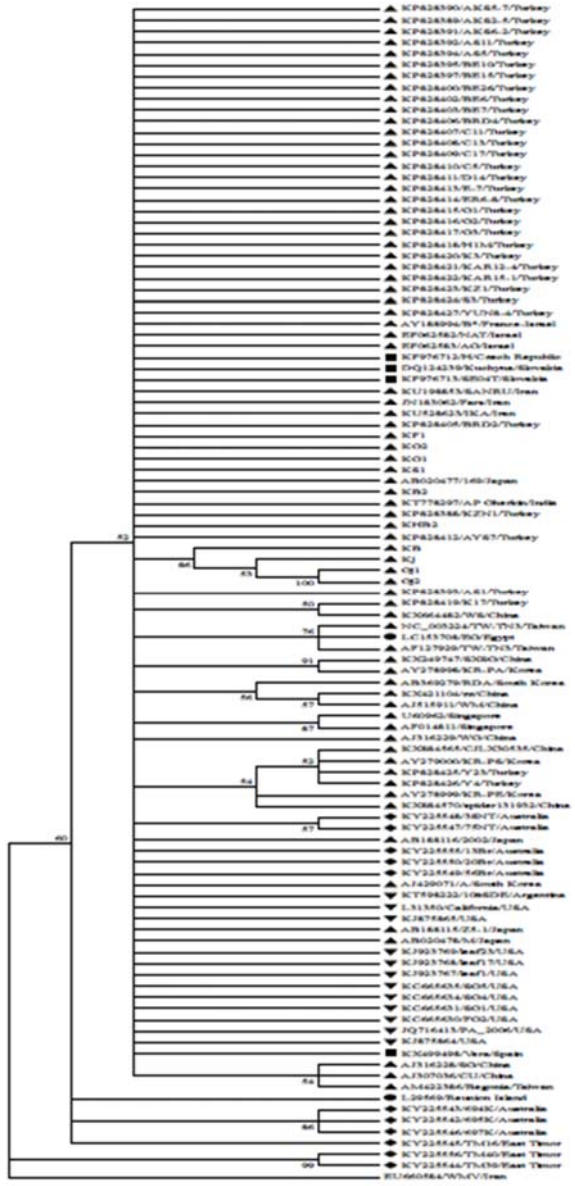

Figure S1. Maximum likelihood phylogenetic tree illustrating the phylogenetic relationships between Iranian and other ZYMV isolates. Tree was drawn by MEGA7 using the CI amino acid sequence. WMV (*Watermelon mosaic virus*) included as out-group. The GenBank accession number, the name of each isolate and its country of origin are listed. Numbers at each node indicate bootstrap percentages based on 1000 replications. Values are shown only when the values are equal or greater than 50%.