Eur J Plant Pathol (2020) 156:1091–1099
https://doi.org/10.1007/s10658-020-01965-3

Molecular variability of watermelon mosaic virus isolates from Argentina

E. Pozzi · M. C. Perotto · S. Bertin · A. Manglli · C. Luciani · V. C. Conci · L. Tomassoli

Accepted: 26 February 2020 / Published online: 12 March 2020

Abstract Watermelon mosaic virus (WMV) is an economically important virus of cucurbit crops in Argentina. The available information on genetic variability must be continuously updated. In this study, we assessed the molecular variability of WMV isolates in Argentina based on the partial sequences of the NIb-CP region and compared them with isolates previously reported from around the world. Forty-six WMV isolates were obtained from naturally infected cucurbit crops collected from 10 provinces between 2011 and 2018. At the molecular level, WMV isolates were grouped into three distinct major phylogenetic groups based on genetic distance. Majority of the Argentine isolates belonged to the emerging group (G3), whereas one isolate was included in G1 and another cluster in G2. G3 was further divided into 5 subgroups, named EM1, EM2, EM3, EM4, and a new group "EMArg", composed exclusively of Argentine isolates with high bootstrap support and high level of significance. Tajima's D and Fu were significantly negative for G3 and EMArg, indicating that the population of this subgroup has recently expanded. AMOVA analysis showed that the isolates was not well correlated with their geographic origin. The fixation index (F_{ST}) value between the WMV isolates from the different provinces in Argentina was lower than 0.33, indicating a relatively frequent gene flow between provinces. In every group, the predominant evolutionary pressure was negative with a mean dN/dS < 1, suggesting a slow replacement fixation rate. One recombination event was detected involving isolates of EM2 cluster from Argentina.

Keywords WMV · Genetic diversity · Cucurbitaceae · Recombination · Selection pressure

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10658-020-01965-3) contains supplementary material, which is available to authorized users.

E. Pozzi · M. C. Perotto · C. Luciani · V. C. Conci
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), CCT, X5020ICA Córdoba, Argentina

M. C. Perotto (✉) · C. Luciani · V. C. Conci
Instituto de Patología Vegetal (IPAVE-CIAP-INTA), Camino 60 cuadras km 5,5, X5020ICA Córdoba, Argentina
e-mail: perotto.cecilia@inta.gob.ar

S. Bertin · A. Manglli · L. Tomassoli
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA), Centro di ricerca Difesa e Certificazione (CREA-DC), Roma, Italy

Introduction

Major cucurbit species, such as squash and pumpkins (*Cucurbita pepo*, *C. moschata*, and *C. maxima*), melon (*Cucumis melo*), cucumber (*Cucumis sativus*), and watermelon (*Citrullus lanatus*), are among the main vegetable crops cultivated in the world. In Argentina, squash, pumpkin, watermelon and melon are widely grown on 32,749 ha (FAOSTAT 2014), and are relevant for the agricultural economy.

Viral diseases of cucurbits cause significant yield losses around the world (Desbiez et al. 2019; Lecoq and Desbiez 2012; Lecoq et al. 1998; Zitter et al. 1996). More than 70 virus species transmitted by several insect vectors are known to infect cucurbit crops worldwide (Lecoq and Katis 2014). Among these, *Watermelon mosaic virus* (WMV), *Zucchini yellow mosaic virus* (ZYMV), and *Papaya ringspot virus* (PRSV), belonging to the genus *Potyvirus*, and *Cucumber mosaic virus* (CMV), genus *Cucumovirus*, are the most commonly reported aphid-borne viruses infecting cucurbits worldwide (Lecoq et al. 1998, Lecoq 2003; Perotto et al. 2016; Vučurović et al. 2012). These viruses are efficiently transmitted in a non-persistent manner by many different aphid species. Watermelon mosaic virus is responsible for economically important losses in Argentina, where its presence was confirmed in all surveys conducted since its first detection in 1971 (Nome et al. 1974). A recent survey carried out in northern and central Argentina showed high values of WMV incidence in several cucurbit species (98% of symptomatic samples collected in 2012 and 100% in 2014; Perotto et al. 2016). Further surveys carried out in 2017 and 2018 confirmed that WMV was the most prevalent and widespread virus found in all cucurbit-growing regions in Argentina, with a relative incidence of 46%, followed by ZYMV (24%) and PRSV (20%) (unpublished data).

Molecular characterization of WMV is based on the nuclear inclusion b (NIb) protein - coat protein (CP) (NIb-CP) region, showing an important genetic variability, with three distinct molecular groups: group 1 (G1), group 2 (G2), and group 3 (G3) (Desbiez et al. 2007). Subsequently, G1 was called classic (CL), whereas isolates belonging to G3 were called emerging (EM), since they correspond to more recent isolates emerged in Europe (only identified since 1999) that progressively replaced CL isolates. The emerging (G3) isolates cause very severe symptoms and are divided into four subgroups, namely EM1 to EM4 (Desbiez et al. 2009).

Given the rapid changes observed in this virus, the information available on genetic variability needs to be continuously updated for a correct interpretation of its complexity. In this study, we described for the first time the population structure of WMV in one country of South America and identified a new EM subgroup composed of Argentine isolates based on the partial sequences of the NIb-CP region.

Materials and methods

Virus isolates

Forty-six WMV isolates were obtained from naturally infected squash, pumpkin, zucchini and melon plants, collected from several locations in Argentina between 2011 and 2018. The survey was conducted the main cucurbit-cultivation areas in the country (Fig. 1), covering 10 provinces: Salta (ST), Mendoza (MZA), San Juan (SJ), La Rioja (LR), Tucumán (TUC), Santiago del Estero (SDE), Córdoba (CBA), Jujuy (JUJ), Buenos Aires (BA), and Santa Fe (StaFe). Samples were properly labelled with the sampling date, crop species, date of planting, growing conditions and foliar symptoms resembling virus infection, and stored at −80 °C in sealed plastic bags until lyophilization for later use (Table 1).

The leaf samples were tested for the presence of WMV by double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), as describe by Clark and Adams (1977), using polyclonal antibodies from Bioreba AG (Switzerland). All the samples that were positive for the presence WMV by DAS-ELISA were then subjected to molecular analysis for genetic characterization.

Molecular analysis

Total RNA was extracted using the Real Total RNA from Tissue and Cell kit (Real, Valencia, Spain) from lyophilized leaf material of WMV infected cucurbit plants following the manufacturer's instructions. The RNA extracts were then analyzed by reverse-transcriptase polymerase chain reaction (RT-PCR). A primer pair was designed for this work, using the Primer3 software (Untergasser et al. 2012) to amplify a 943-bp fragment across the NIb-CP region: WMV NIb-CP 5′ (CGTGCTGTAARCAAGGKTGGTC) and WMV NIb-CP 3′ (CCATTCATWATCACACCCATTTG MTC). One-step RT-PCR protocol was performed in a total volume of 25 μl containing 2 μl of total RNA extract, 1X GoTaq® Reaction Buffer (Promega), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.4 μM of each primer, 0.75 U GoTaq®G2 DNA Polymerase (Promega), 1.2 U AMV Reverse Transcriptase (Promega) and 20 U RNaseOUT™ Recombinant Ribonuclease Inhibitor (Invitrogen, ThermoFisher Scientific). Amplification was performed according to the

Fig. 1 Map of Argentina indicating (in grey) the provinces included in this survey. Note: ST: Salta, MZA: Mendoza, SJ: San Juan, LR: La Rioja, TUC: Tucumán, SDE: Santiago del Estero, CBA: Córdoba, JUJ: Jujuy, BA: Buenos Aires, StaFe: Santa Fe

following conditions: reverse transcription at 46 °C for 30 min, followed by denaturation at 95 °C for 5 min, and by 34 cycles of the following steps: 30 s at 94 °C, 30 s at 56 °C annealing and 45 s at 72 °C with a final extension at 72 °C for 10 min. The RT-PCR products were sequenced (EUROFINS GENOMICS, Milan) in both directions and the assembled sequences were deposited in GenBank (NCBI accession numbers in Table 1).

Sequence analysis

The NIb-CP fragments were aligned using the Muscle and MAFF programs included in Geneious R9 software (Kearse et al. 2012). The average pairwise distance between and within groups (G1, G2 and G3) and

Table 1 Watermelon mosaic virus isolates collected and sequenced from 10 different provinces in Argentina from the 2011 to 2018

Acc. number	Growing season	Province	Host species
MN006911	2011/2012	SDE	*Cucumis melo*
MN006932	2011/2012	SJ	*Cucumis melo*
MN006934	2011/2012	SJ	*Cucumis melo*
MN019129	2012/2013	SJ	*Cucumis melo*
MN006909	2015/2016	SJ	*Cucumis melo*
MN006925	2015/2016	SJ	*Cucumis melo*
MN006937	2015/2016	SJ	*Cucumis melo*
MN019130	2015/2016	SJ	*Cucumis melo*
MN006945	2016/2017	SJ	*Cucumis melo*
MN006946	2016/2017	SJ	*Cucumis melo*
MN006913	2017/2018	CBA	*Cucumis melo*
MN006907	2017/2018	SDE	*Cucumis melo*
MN006914	2017/2018	LR	*Cucurbita lagenaria*
MN006915	2017/2018	LR	*Cucurbita lagenaria*
MN006939	2011/2012	SDE	*Cucurbita maxima*
MN006940	2011/2012	SDE	*Cucurbita maxima*
MN006941	2011/2012	SDE	*Cucurbita maxima*
MN006942	2012/2013	BA	*Cucurbita maxima*
MN019131	2015/2016	BA	*Cucurbita maxima*
MN006912	2015/2016	ST	*Cucurbita maxima*
MN006921	2016/2017	StaFe	*Cucurbita maxima*
MN006904	2017/2018	CBA	*Cucurbita maxima*
MN006905	2017/2018	CBA	*Cucurbita maxima*
MN006928	2017/2018	JUJ	*Cucurbita maxima*
MN006929	2017/2018	JUJ	*Cucurbita maxima*
MN006930	2017/2018	JUJ	*Cucurbita maxima*
MN006938	2017/2018	TUC	*Cucurbita maxima*
MN006935	2016/2017	CBA	*Cucurbita maxima* subs. *andreana*
MN006916	2015/2016	MZA	*Cucurbita moschata*
MN006944	2015/2016	SJ	*Cucurbita moschata*
MN006920	2015/2016	ST	*Cucurbita moschata*
MN006931	2016/2017	StaFe	*Cucurbita moschata*
MN006919	2017/2018	MZA	*Cucurbita moschata*
MN006922	2017/2018	MZA	*Cucurbita moschata*
MN006908	2017/2018	SDE	*Cucurbita moschata*
MN006910	2017/2018	SDE	*Cucurbita moschata*
MN006917	2015/2016	MZA	*Cucurbita pepo*
MN006918	2015/2016	MZA	*Cucurbita pepo*
MN006943	2015/2016	MZA	*Cucurbita pepo*
MN006927	2015/2016	SJ	*Cucurbita pepo*
MN006923	2015/2016	ST	*Cucurbita pepo*
MN006924	2015/2016	ST	*Cucurbita pepo*
MN006926	2015/2016	ST	*Cucurbita pepo*
MN006933	2016/2017	MZA	*Cucurbita pepo*
MN006936	2016/2017	MZA	*Cucurbita pepo*
MN006906	2017/2018	TUC	*Cucurbita* sp.

ST Salta, *MZA* Mendoza, *SJ* San Juan, *LR* La Rioja, *TUC* Tucumán, *SDE* Santiago del Estero, *CBA* Córdoba, *JUJ* Jujuy, *BA* Buenos Aires, *StaFe* Santa Fe

subgroups (EM1, EM2, EM3, EM4, EMArg) was calculated using *p*-distances with standard error. A maximum-likelihood (ML) consensus tree was generated from 1000 bootstrap replicates. Best-fit substitution models were selected according to BIC and AIC. The Tamura Nei model with discrete Gamma distribution was the best-fit model (Tamura and Nei 1993; Kumar et al. 2016). Maximum Likelihood (ML), Neighbor Joining (NJ), and Maximum Parsimony (MP) trees were also generated. All this sequence analysis were performed by MEGA 7 (Kumar et al. 2016).

Population genetics

Tajima's D-test and Fu's test for haploid sequences (Tajima 1989; Fu 1997) were performed using Arlequin 3.5 software (Excoffier and Lischer 2010) to compute the nucleotide diversity of Argentine isolates within each group and subgroup. These tests compare the number of low and intermediate frequency mutations in an alignment, and detect natural selection in gene sequences.

The extent of genetic differentiation and the levels of gene flow between populations (provinces) were estimated via the Fixation index (Fst) using Arlequin 3.5. The absolute value of Fst ranges from 0 (undifferentiated populations) to 1 (fully differentiated populations). An absolute value of Fst > 0.33 suggests an infrequent gene flow, whereas an absolute value of Fst < 0.33 suggests a frequent gene flow between populations (Rozas et al. 2003, 2017).

The distribution of the genetic variability among Argentine samples was tested via the analysis of molecular variance (AMOVA) implemented in Arlequin 3.5, with provinces being considered populations. The significance of the analyses was obtained by performing 1023 permutations.

Selection analysis

Selection pressure was measured by estimating the ratio of non-synonymous (dN) to synonymous (dS) substitution rates .

Site-specific selection events were identified by including single likelihood ancestor counting (SLAC), random effects likelihood (REL), and branch site unrestricted statistical test for episodic diversification (BUSTED). All methods are available at the Datamonkey online server (http://www.datamonkey.org) (Weaver et al. 2018).

Recombination analysis

Recombination events were first searched visually with SimPlot (Ray 1998), using the NJ model and setting 1000 bootstrap replicates and the parental threshold of 70. The analysis was performed on the aligned sequences (window 200 bp, step 20 bp), with a Kimura (2-parameter) correction for multiple substitutions.

Evidence of recombination was further assessed using different algorithms compiled in the RDP4 (Martin et al. 2015): RDP, GENECONV, Chimaera, MaxChi, BOOTSCAN, SISCAN, LARD, PhylPro and 3Seq, with a Bonferroni-corrected *p* value cutoff of 0.05.

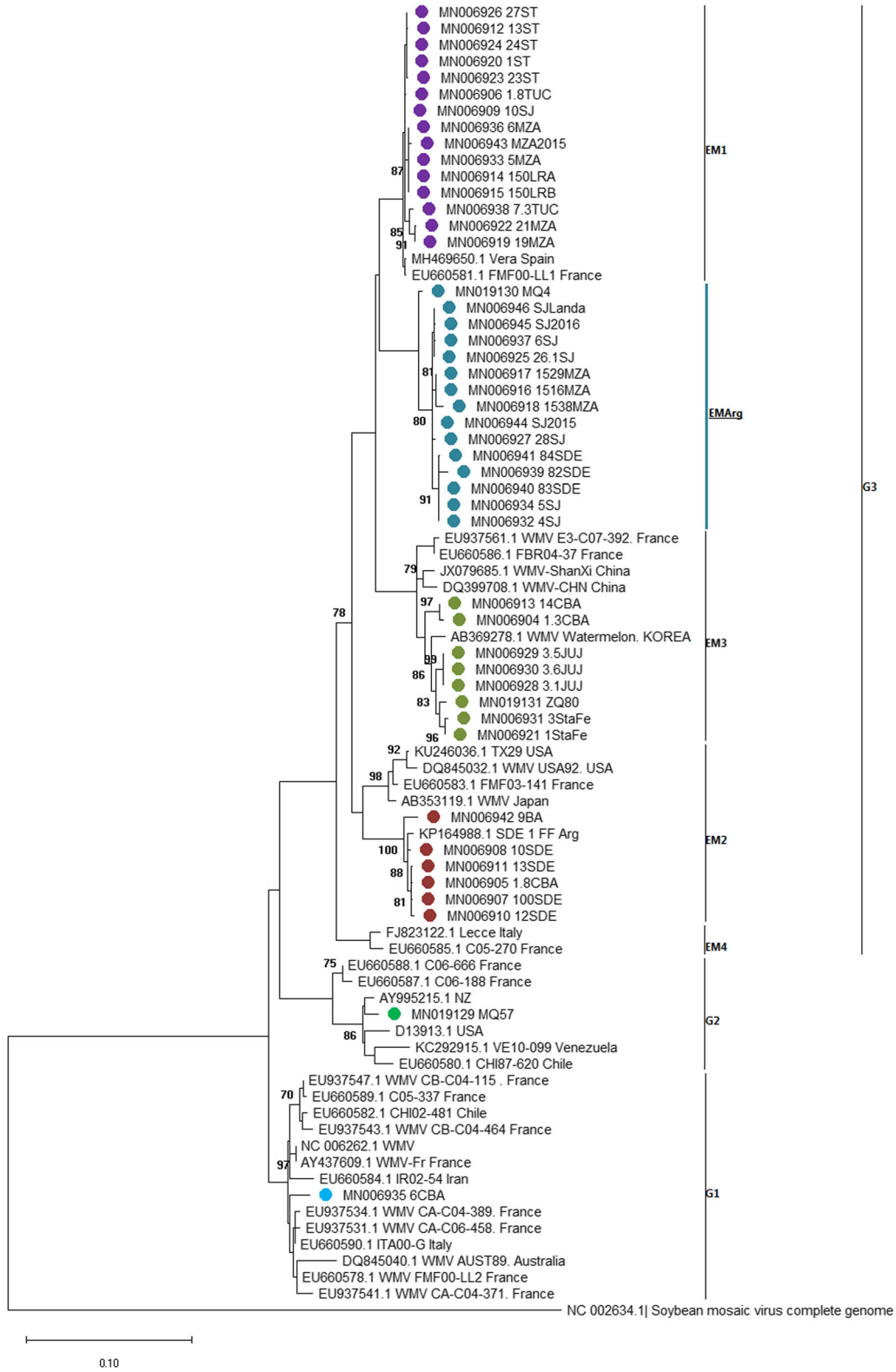
Results

Sequence analysis

A 943-bp NIb-CP fragment was obtained for all the 46 Argentine WMV isolates collected from 2011 to 2018. From nucleotide sequences of these isolates and others available in GenBank, an ML phylogenetic tree was obtained (Fig. 2). The NJ and MP trees constructed produced the same result (data not shown). The WMV isolates were grouped into 3 distinct major clades representing the genetic groups G1 to G3. Among the Argentine samples, the isolates 6CBA (MN006935) clustered in G1 and MQ57 in G2 (MN019129), whereas the remaining 44 isolates were included in G3. The protein translation confirms the arrangement of the isolates in the phylogenetic tree. Within G3, the Argentine samples were distributed in the subgroups EM1, EM2 and EM3. Moreover, 15 samples from SJ, MZA and SDE provinces formed a separated subgroup, hereafter referred to as "EMArg".

Intragroup average *p*-distance of the NIb-CP sequences was higher in G3 (0.045 ± 0.012) than in G1 and G2 ($p = 0.031 \pm 0.009$ and $p = 0.015 \pm 0.007$, respectively). Inter-

Fig. 2 Phylogenetic tree obtained by Maximum Likelihood method based on the Tamura-Nei model. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The tree is displayed with the *Soybean mosaic virus* as an outgroup. Bootstrap values (1000 bootstraps) above 70% are indicated for each node. The Argentinian isolates obtained in this study are identified with a full circle


MN006926 27ST
MN006912 13ST
MN006924 24ST
MN006920 1ST
MN006923 23ST
MN006906 1.8TUC
MN006909 10SJ
MN006936 6MZA
MN006943 MZA2015
MN006933 5MZA
MN006914 150LRA
MN006915 150LRB
MN006938 7.3TUC
MN006922 21MZA
MN006919 19MZA
MH469650.1 Vera Spain
EU660581.1 FMF00-LL1 France
MN019130 MQ4
MN006946 SJLanda
MN006945 SJ2016
MN006937 6SJ
MN006925 26.1SJ
MN006917 1529MZA
MN006916 1516MZA
MN006918 1538MZA
MN006944 SJ2015
MN006927 28SJ
MN006941 84SDE
MN006939 82SDE
MN006940 83SDE
MN006934 5SJ
MN006932 4SJ
EU937561.1 WMV E3-C07-392. France
EU660586.1 FBR04-37 France
JX079685.1 WMV-ShanXi China
DQ399708.1 WMV-CHN China
MN006913 14CBA
MN006904 1.3CBA
AB369278.1 WMV Watermelon. KOREA
MN006929 3.5JUJ
MN006930 3.6JUJ
MN006928 3.1JUJ
MN019131 ZQ80
MN006931 3StaFe
MN006921 1StaFe
KU246036.1 TX29 USA
DQ845032.1 WMV USA92. USA
EU660583.1 FMF03-141 France
AB353119.1 WMV Japan
MN006942 9BA
KP164988.1 SDE 1 FF Arg
MN006908 10SDE
MN006911 13SDE
MN006905 1.8CBA
MN006907 100SDE
MN006910 12SDE
FJ823122.1 Lecce Italy
EU660585.1 C05-270 France
EU660588.1 C06-666 France
EU660587.1 C06-188 France
AY995215.1 NZ
MN019129 MQ57
D13913.1 USA
KC292915.1 VE10-099 Venezuela
EU660580.1 CHI87-620 Chile
EU937547.1 WMV CB-C04-115 . France
EU660589.1 C05-337 France
EU660582.1 CHI02-481 Chile
EU937543.1 WMV CB-C04-464 France
NC 006262.1 WMV
AY437609.1 WMV-Fr France
EU660584.1 IR02-54 Iran
MN006935 6CBA
EU937534.1 WMV CA-C04-389. France
EU937531.1 WMV CA-C06-458. France
EU660590.1 ITA00-G Italy
DQ845040.1 WMV AUST89. Australia
EU660578.1 WMV FMF00-LL2 France
EU937541.1 WMV CA-C04-371. France
NC 002634.1| Soybean mosaic virus complete genome
EM1
EMArg
EM3
EM2
EM4
G2
G1
G3
87
85
91
81
80
91
79
78
97
99
86
83
96
92
98
100
88
81
75
86
70
97
0.10

group variability was higher between G1 and G3 ($p = 0.089 \pm 0.022$) than between G1 and G2, and between G2 and G3 ($p = 0.057 \pm 0.019$ and $p = 0.073 \pm 0.021$, respectively). Considering only the Argentine G3 sequences, the EMArg with EM1 isolates, showed one of the lowest intra-subgroup variabilities $p = 0.007 \pm 0.001$ and $p = 0.005 \pm 0.001$ respectively, followed by EM2 ($p = 0.008 \pm 0.002$) and EM3 ($p = 0.013 \pm 0.003$). The genetic distance of EMArg isolates was 0.05 (± 0.008) from EM1 and EM2 subgroups, and 0.07 (± 0.009) from EM3.

Population genetics and selection analysis

Tajima's D test and Fu's test, performed on the Argentine sequences comprised in the emerging group (G3) were negative and statistically significant (D` = −2.03481, $P = 0.002$; FS = −12.63464, $P = 0.003$). Moreover, the new subgroup EMArg was also significantly negative for the Tajima's D test (D` = −2.43774, $P = 0.000$).

The AMOVA analyses showed that most of the total genetic variability of the Argentine isolates is distributed within each province (76.98%) and not among them (23.02%) ($P < 0.05$, 1023 permutations). The Fst value between the WMV isolates was 0.23024, suggesting a frequent gene flow between provinces.

The dN/dS ratio was investigated to assess the presence of selective pressures influencing the population composition of WMV in Argentina. No evidence for positive selected sites was found within the NIb-CP sequences using the codon-based SLAC, REL and BUSTED methods. The absence of positive selection was confirmed by the low value of the overall dN/dS ratio (0.125) estimated for the Argentine isolates.

Recombination analysis

A recombination event was detected visually with SimPlot and then confirmed with RDP4 analysis. The event was detected by three different methods, MaxChi, Chimaera and 3Seq, with values of $P < 1 \times 10^{-3}$ for the isolates clustering in EM2. A recombinant fragment of 389 nt was identified between the positions 119 and 508 of the NIb terminal region. The putative parents were MN006939 82SDE (major parent) and MN006922 21MZA (minor parent), which are both from Argentina and belong to the EMArg and EM1 subgroups, respectively (Table 2).

Table 2 Summary of the recombination event detected at *NIb-CP* genomic regions of 46 isolates of Watermelon mosaic virus (WMV) from Argentina

Recombinant isolate (s)	Putative		Score for the nine detection methods in RDP4								
	Minor parental	Mayor parental	RDP	GENECOV	Chimaera	MaxChi	BOOTSCAN	SISCAN	LARD	PhylPro	3Seq
MN006942 9BA	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
KP164988 SDE 1FF	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
MN006908 10SDE	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
MN006911 13SDE	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
MN006905 1.8CBA	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
MN006907 100SDE	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03
MN006910 12SDE	MN006922 21MZA	MN006939 82SDE	NS	NS	7.79E-03	2.34E-03	NS	NS	NS	NS	2.79E-03

The event was identified by at least three different methods implemented in the RDP4 v.4.97 program, with an associated $P < 10^{-3}$

Discussion

Watermelon mosaic virus is one of the most prevalent potyviruses causing severe damages in cucurbits worldwide, with Argentina being no exception. The three distinct WMV molecular groups are known to show different origins and symptoms. The isolates belonging to G1 are considered "classical" isolates due to their early appearance in the Mediterranean basin, and are associated with mild symptoms. G2 represents a distinct molecular lineage and this group originated from different parts of the world. The emerging isolates belonging to G3 appeared in Europe early in the 2000s; they are associated with an increase in symptom severity and rapidly replaced the CL isolates (Desbiez et al. 2007, 2009; Wang et al. 2017).

In a previous study a complete sequence of an Argentine WMV collected in 2012 was obtained, which was found to belong to G3 (Perotto et al. 2016). In this study, most of the Argentine isolates were located in the emerging groups (G3). Only two isolates from MZA clustered in G1 (classic) and G2 (Fig. 2). On the other hand, the three groups correlated with an amino acid motif at the N-terminal extremity of the CP. G1 usually had a "KEA" motif, whereas G2 displayed a "KET", and G3 had a "KEKET" (Desbiez et al. 2007).

The WMV-EM isolates are known to be more virulent than WMV-CL isolates and the virus accumulation is higher than for CL isolates (Juarez et al. 2013). These factors favored a rapid expansion of EM isolates in the Mediterranean basin and almost led to the disappearance of the CL strain in this area (Lecoq and Desbiez 2012; Joannon et al. 2010). Information on the temporal genetic evolution of WMV populations in Argentina is scarce, since only one old Argentine isolate collected in 1994 (accession number DQ845043) belonging to G1 has been studied and deposited (Desbiez et al. 2007). Therefore, it is difficult to state that this group was sporadic in Argentina or if a replacement of CL by EM isolates has occurred in the recent past, as reported in several Mediterranean countries.

In the phylogenetic tree, the clusters corresponding to the main groups G1–3 were genetically homogeneous and mainly not fragmented. On the other hand, the intergroup variability was greater than the intragroup variability, as reported in other studies (Desbiez et al. 2009). These results as well as the lack of historical genetic information prevent us from inferring whether the Argentine emerging strains originated from new introductions or from evolution of local populations. In the phylogenetic tree, the separation of a new Argentine subgroup (EMArg), different from the already described for the European isolates (Desbiez et al. 2007), can be observed, which is supported by high bootstrap values. The statistical tests evaluated using Tajima's D and Fu were significantly negative for G3 and EMArg, indicating the occurrence of the recent expansion of the population of the subgroup EMArg.

Comparison of nucleotide sequence data from plant virus isolates of different geographic origins has allowed the analysis of the genetic structure for some plant virus species across most of their geographic range (Moreno et al. 2004). In this study, AMOVA analysis showed that most of the genetic variability of the Argentine samples was distributed within rather than among provinces, indicating that the molecular clustering of isolates was not well correlated with their geographic origin. The absence of spatial differentiation of the populations is consistent with the observations made in France and Spain by Desbiez et al. (2007) and Moreno et al. (2004), respectively. Similarly, the Fst value between the WMV isolates from the different provinces in Argentina was lower than 0.33, indicating a relatively frequent gene flow between provinces.

The average value of the dN/dS ratio for the Argentine isolates was <1, indicating that the genetic selection was mostly against amino acid changes, namely, "negative selection" or "purifying selection". The SLAC, REL and BUSTED analyses performed site by site confirm the absence of positive selection events. The power of the dN/dS statistic is known to be reduced when samples are isolates within a virus species, since the detected polymorphisms might not be fixed in divergent populations (Kryazhimskiy and Plotkin 2008). A side effect is that the ratio values for those genes evolving under negative selection are typically closer to one when intra-specific rather than inter-specific samples are compared. The dN/dS value obtained in this study was much lower than one, indicating that no negative selection events were missed. A second issue is that the power of dN/dS statistic might fail to detect positive selection of close samples. However, neither dN/dS values nor site per site analyses identified any positively selected sites within our sequences. These results are consistent with the previous studies based on WMV full-sequences (Desbiez and Lecoq 2008) as well as with the general knowledge that RNA plant viruses commonly undergo to negative rather than

positive selection (Hughes and Hughes 2007). Therefore, even if we cannot exclude that positive selection events have been underestimated in our analysis, we still consider the absence of diversifying processes as a reasonable hypothesis.

Recombination events represent an important source of genetic variation for potyviruses (Gibbs and Ohshima 2010). In this study, one recombinant event was detected, which included seven of the 46 WMV isolates. Interestingly, the recombination event involves only the Argentine isolates belonging to the EM2 cluster. No recombination signals were detected for the other EM2 isolates from USA, France or Japan. The putative parental sequences were from two different provinces, one from SDE (major parent) and the other from MZA (minor parent). This finding supports the gene flow between different geographical origins.

To the best of our knowledge, this is the first study of the population structure of WMV in South America. Combined management strategies that involve crop rotation, planting dates, weed control, and resistant cultivars might contribute to prevention of WMV spread. However, further research is needed to understand the rapid and continuous changes of this virus.

Acknowledgements This study was carried out at CREA-DC (Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Centro di ricerca Difesa e Certificazione) in Rome, Italy, and at IPAVE-INTA (Instituto de Patología Vegetal-Instituto Nacional de Tecnología Agropecuaria) in Argentina. This study was partially supported by CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) and a scholarship from IILA-Organizzazione Internazionale Italo-Latino Americana.

Compliance with ethical standards

Conflict of interest All authors declare no conflict of interest.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

References

Clark, M. F., & Adams, A. N. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology, 34*(3), 475–483.

Desbiez, C., & Lecoq, H. (2008). Evidence for multiple intraspecific recombinants in natural populations of watermelon mosaic virus (WMV, Potyvirus). *Archives of Virology*. https://doi.org/10.1007/s00705-008-0170-2.

Desbiez, C., Costa, C., Wipf-Scheibel, C., Girard, M., & Lecoq, H. (2007). Serological and molecular variability of watermelon mosaic virus (genus Potyvirus). *Archives of Virology, 152*(4), 775–781.

Desbiez, C., Joannon, B., Wipf-Scheibel, C., Chandeysson, C., & Lecoq, H. (2009). Emergence of new strains of watermelon mosaic virus in South-Eastern France: evidence for limited spread but rapid local population shift. *Virus Research, 141*(2), 201–208.

Desbiez, C., Verdin, E., Moury, B., Lecoq, H., Millot, P., Wipf-Scheibel, C., Mirzayeva, S., Sultaova, N., Balakishiyeva, G., Mammadov, A., Kheyr-Pour, A., & Huseynova, I. (2019). Prevalence and molecular diversity of the main viruses infecting cucurbit and solanaceous crops in Azerbaijan. *European Journal of Plant Pathology, 153*, 359–369. https://doi.org/10.1007/s10658-018-1562-0.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources, 10*(3), 564–567.

FAOSTAT (2014). Food and agriculture Organization of the United Nations. Statistics Division http://www.fao.org/faostat/en/#compare. Accessed 03 Jan 2019.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*(2), 915–925.

Gibbs, A., & Ohshima, K. (2010). Potyviruses and the digital revolution. *Annual Review of Phytopathology*. https://doi.org/10.1146/annurev-phyto-073009-114404.

Hughes, A. L., & Hughes, M. A. (2007). More effective purifying selection in RNA viruses than in DNA viruses. *Gene, 404*(1–2), 117–125.

Joannon, B., Lavigne, C., Lecoq, H., & Desbiez, C. (2010). Barriers to gene flow between emerging populations of watermelon mosaic virus in southeastern France. *Phytopathology, 100*(12), 1373–1379.

Juarez, M., Legua, P., Mengual, C. M., Kassem, M. A., Sempere, R. N., Gómez, P., & Aranda, M. A. (2013). Relative incidence, spatial distribution and genetic diversity of cucurbit viruses in eastern Spain. *Annals of Applied Biology, 162*(3), 362–370.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bts199.

Kryazhimskiy, S., & Plotkin, J. B. (2008). The population genetics of dN/dS. *PLoS Genetics, 4*(12), e1000304.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution, 33*(7), 1870–1874.

Lecoq, H. (2003). Cucurbits. In G. Loebenstein & G. Thottappilly (Eds.), *Virus and virus-like diseases of major crops in developing countries* (pp. 665–687). New York. ISBN-13: 9781402012303: Springer.

Lecoq, H., & Desbiez, C. (2012). Viruses of cucurbit crops in the Mediterranean region: an ever-changing picture. *Advances in Virus Research*. https://doi.org/10.1016/B978-0-12-394314-9.00003-8.

Lecoq, H., & Katis, N. (2014). Control of cucurbit viruses. *Advances in Virus Research*. https://doi.org/10.1016/B978-0-12-801246-8.00005-6.

Lecoq, H., Wisler, G., & Pitrat, M. (1998). Cucurbit viruses: The classics and the emerging. In J. D. McCreight (Ed.), *Cucurbitaceae '98, evaluation and enhancement of cucurbit germplasm* (pp. 126–142). Alexandria: ASHS.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*. https://doi.org/10.1093/ve/vev003.

Moreno, I. M., Malpica, J. M., Dıaz-Pendón, J. A., Moriones, E., Fraile, A., & Garcıa-Arenal, F. (2004). Variability and genetic structure of the population of watermelon mosaic virus infecting melon in Spain. *Virology, 318*(1), 451–460.

Nome, S. F., March, G. J., Giorda, L. M. (1974). Disminución de la productividad de plantas de zapallito de tronco (Cucurbita máxima Duch var. zapallito (Carr. Millán) infectadas con el virus del mosaico de la sandía, raza 2 (Watermelon Mosaic Virus-2). *IDIA*, 321–324.

Perotto, M. C., Celli, M. G., Pozzi, E. A., Luciani, C. E., & Conci, V. C. (2016). Occurrence and characterization of a severe isolate of watermelon mosaic virus from Argentina. *European Journal of Plant Pathology, 146*(1), 213–218.

Ray, S. C. (1998). SimPlot for windows (version 1.6). Baltimore, Md. Distributed by author. http://www.welch.jhu.edu/sray/download

Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btg359.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msx248.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585–595.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution, 10*(3), 512–526.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3–new capabilities and interfaces. *Nucleic Acids Research, 40*(15), e115.

Vučurović, A., Bulajić, A., Stanković, I., Ristić, D., Berenji, J., Jović, J., & Krstić, B. (2012). Non-persistently aphid-borne viruses infecting pumpkin and squash in Serbia and partial characterization of Zucchini yellow mosaic virus isolates. *European Journal of Plant Pathology, 133*(4), 935–947.

Wang, D., Li, G., & Du, S. S. (2017). Occurrence of viruses infecting melon in Xinjiang of China and molecular characterization of watermelon mosaic virus isolates. *European Journal of Plant Pathology, 147*, 919–931. https://doi.org/10.1007/s10658-016-1060-1.

Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., Kosakovsky Pond, S. L. (2018). *Molecular Biology and Evolution*, (Volume 35, Issue 3, Pages 773–777). Internet Resource: http://www.datamonkey.org

Zitter, T. A., Hopkins, D. L., & Thomas, C. E. (Eds.). (1996). *Compendium of cucurbit diseases* (87 p). St. Paul: APS Press.