Australasian Plant Pathol. (2016) 45:401–409
DOI 10.1007/s13313-016-0426-3

ORIGINAL PAPER

Genetic variability and recombination analysis of the coat protein gene of *Strawberry mild yellow edge virus*

A. K. Torrico[1] · M. G. Celli[2] · E. E. Cafrune[1] · D. S. Kirschbaum[3] · V. C. Conci[1,2]

Received: 12 August 2015 / Accepted: 13 June 2016 / Published online: 6 July 2016
© Australasian Plant Pathology Society Inc. 2016

Abstract *Strawberry mild yellow edge virus* (SMYEV) has been detected in most of the strawberry production regions worldwide. However, little is known about differences between distinct isolates. The aim of this study was to enhance the knowledge about the genetic variability of different SMYEV isolates, exploring the phylogenetic relationships and assessing recombinant events among them. The coat protein (CP) gene of 12 Argentinian SMYEV isolates was sequenced. There were 729 nucleotides (nt) in all of the isolates, encoding a protein of 242 amino acids (aa). Argentina isolates shared 81.5–99.6 % nucleotide identity. The comparison of these isolates with 30 SMYEV sequences from other countries published in the GenBank, revealed an identity ranging from 81.6 to 99 %. The phylogenetic analysis showed the presence of four possible subgroups, with the Argentinian isolates being included in all of them. Recombination analysis indicated that 16–2 (KP 284155) Argentinian and AJ577342 Chilean isolates are recombinant and that they are a result of recombination events where parts of the genome were exchanged between different SMYEV sequences.

✉ V. C. Conci
conci.vilma@inta.gob.ar

[1] Instituto de Patología Vegetal (IPAVE), Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (INTA), Camino 60 cuadras km 5,5 (5119), Córdoba, Argentina

[2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Córdoba, Argentina

[3] Estación Experimental Agropecuaria Famaillá - Instituto Nacional de Tecnología Agropecuaria (EEAF-INTA), Ruta Provincial 301 km 32 (4132), Famaillá, Tucumán, Argentina

Keywords *Potexvirus* · *Fragaria* x *ananassa* · Phylogenetic analysis · Molecular characterization · SMYEV

Introduction

Aphid-transmitted viruses are the most important and widely studied group of viruses infecting strawberry (*Fragaria* x *ananassa* Duch.) (Tzanetakis and Martin 2013). *Strawberry mottle virus* (SMoV), *Strawberry crinkle virus* (SCV), *Strawberry vein banding virus* (SVBV) and *Strawberry mild yellow edge virus* (SMYEV) are usually found in mixed infections producing severe yield losses in different strawberry producing countries worldwide (Thompson and Jelkmann 2003; Chang et al. 2007). These viruses naturally infect the genus *Fragaria.* Simultaneous infections of the same plant by several viruses in commercial strawberry cultivars have been found to produce dwarfism, interveinal yellowing and small leaves, chlorotic mottle, leaf deformation and important yield losses (Maas 1998). SMYEV was first reported in Argentina in 2008 in single and mixed infections (Conci et al. 2009). This virus, which belongs to the *Potexvirus* genus, (*Alphaflexiviridae* family) has flexuous filamentous particles typical of the genus and a single-stranded positive-sense RNA of 5.9–7 kb in length (King et al. 2012).

Molecular characterization is currently the most widely used tool to classify viral species. Viral species and their strains or variants can be differentiated based on knowledge of their genome sequences. The size of the genome fragment that is necessary to be compared for this classification is a matter of debate. Species of the *Potexvirus* genus are differentiated by comparing the RNA-dependent RNA polymerase and complete coat protein (CP) sequences. The Ninth Report of the International Committee on Taxonomy of Viruses (ICTV) states that distinct species of *Potexvirus* have less than

~72 % identical nucleotides (nt) or 80 % identical amino acids (aa) between their CP or polymerase genes (King et al. 2012).

Variability is an intrinsic characteristic of systems with reproductive capacity, determining the appearance of individuals that differ genetically from their parents (García Arenal et al. 2008). Genetic variation is generated by changes occurring during the replication of virus genomes, with mutation and recombination being the most important ones. Mutation is an inheritable change in base sequences of nucleic acids contained in the genome of an organism, whereas recombination is the process involving exchange of genetic information segments between the nucleotide chains of different genetic variants during replication. As a consequence, recombination is the result of a genetic exchange (García Arenal et al. 2001). Variability of SMYEV in nature is poorly known. A previous work, that conducted phylogenetic analyses of different sequences of the CP gene of this virus, found three different groups (Thompson and Jelkmann 2004).

The phylogenetic analysis is based on relationships of evolutionary proximity among different species, thereby reconstructing the history of their diversification (phylogenesis). It reveals the global situation and does not necessarily reflect the frequency of RNA recombination itself. The exchange of genetic material occurs more frequently within a viral population, although it also happens among different viral strains or among different viruses. In addition, it has been shown that viral RNAs might recombine with the host's RNA, as well as with the viral RNA expressed by transgenic hosts (García Arenal et al. 2001).

The present work reports the gene sequences coding for the CP of different Argentinian SMYEV isolates. A phylogenetic comparison with other isolates from the different parts of the world, and a recombination analysis among them are also discussed.

Materials and methods

Viral sources and coat protein amplification

Strawberry plants, cultivars Camarosa and Albion, both symptomless and with striking virus symptoms (leaf deformation, mosaic, dwarfism and stunting) were collected from different production fields in Lules, Tucumán province, Argentina. The collected plants were maintained in greenhouses under controlled conditions at the *Instituto de Patología Vegetal* (IPAVE-INTA) in Córdoba, Argentina.

Leaf samples were analyzed using DAS-ELISA with antisera (BIOREBA Latin America SA) against SMYEV, *Tobacco necrosis virus*, *Tomato black ring virus*, *Raspberry ringspot virus*, *Arabis mosaic virus*, *Strawberry latent ringspot virus*, *Apple mosaic virus* and *Tomato ringspot virus*, according to the manufacturer's specifications. SMYEV infected samples were used to amplify the entire CP gene of the virus (Table 1).

Total nucleic acid was extracted from strawberry leaves using the modified CTAB protocol described by Chang et al. (2007). The extraction buffer consisted of 2 % CTAB; 100 mm Tris, pH 8.0; 20 mm EDTA; 1.4 m NaCl, with 2 % ß-mercaptoetanol added just prior to use. Leaf material was ground with liquid nitrogen; 0,05 g of powdered tissue from each sample was mixed with 500 μL of extraction buffer and vigorously homogenized. Samples were incubated for 20 min at 65 °C. Afterwards, suspensions were extracted twice with chloroform/isoamyl alcohol (24:1) with the phases separated by centrifugation. The supernatant phase was collected again and transferred to a new tube and then added 1/10 vol. NaAc (pH 5.2) and 2.5 vol. ethanol were added and mixed by inversion. The mix was incubated 40 min at −70 °C. After that, samples were centrifuged at 12,000 rpm at 4 °C during 30 min. The supernatant phase was discarded and the pellet formed in the bottom of 1.5 ml tubes was washed in 70 % ethanol and air-dried before resuspension in 30 μl of nuclease free water. Viral RNA was transcribed to cDNA with M-MLV reverse transcriptase (200 U) (Promega, Madison, WI, USA) according to the manufacturer's instructions and using random primer (N_6) (50 μM) and Oligo $d(t)_{15}$ primer (50 μM) (Biodynamic SRL, Bs As, Argentina). Multiplex PCR was used for detection of SMYEV, SMoV and SVBV with the specific primers YT1/Y2, D1/D3 and I2/SM2, respectively (Sui et al. 2003; Thompson et al. 2003; Thompson and Jelkmann 2004; Chang et al. 2007) carried out as described Chang et al. (2007). The reaction mix consisted of 1 μl of first strand cDNA, 4 μl PCR buffer (5×), 0.2 μl Go Taq DNA polymerase, 5 U/μl (Promega, Madison, WI, USA), a 1.6 μl of dNTP mix (2.5 μmol), 0.35, 0.50 and 0.30 μl of the primers D1/D3(0.17 μM), I2/SM2 (0.25 μM) and YT1/Y2 (0.15 μM), each and 10.9 μl nuclease free water to a final of 20 μl. Multiplex PCR conditions were initial denaturation step at 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 55 °C for 30 s and 68 °C for 30 s, and then a final elongation step at 72 °C for 5 min. The same conditions were used for single PCR reaction for amplification of SMYEV, with exception of quantity primer (0.4 μl) for each reaction.

For detection of SCV, cDNA synthesis was performed as described before. The nested PCR was carried out as described by Posthuma et al. (2002). First round of amplification with SCD1FW and SCD5RV, PCR mix (20 μl total volume) contained 4 μl PCR buffer (5×) 0.5 μl of dNTP mix (10 mM) and 0.2 μl Go Taq DNA polymerase, 5 U/μl (Promega, Madison, WI, USA). The cycling conditions consisted of an initial denaturation at 94 °C for 2 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 54 °C for 30 s and extension at 72 °C for 1 min; and a final extension step at 72 °C for 5 min. One microlitre of this first amplification reaction was used for the nested PCR with SCD2FW and SCD1RV. The

Table 1 Characteristic of SMYEV isolates compared in this study

Name of isolate / accession number	Country	Host	Infection/symptoms	Virus detected	Reference
13–2-5 / KP284152	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SCV	This study
13–3-4 / KP284153	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SCV	This study
15 / KP284154	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SMoV	This study
16–1/ KP284162	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SMoV, SCV	This study
16–2 / KP284155	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SMoV, SCV	This study
16–4-5 / KP284156	Argentina	F. x ananassa cv. Camarosa	Mixed /foliar symptoms	SMYEV, SMoV, SCV	This study
36–1-3 / KP284157	Argentina	F. x ananassa cv. Albión	Single/symptomless	SMYEV	This study
36–2–4-6 / KP284158	Argentina	F. x ananassa cv. Albión	Single/ symptomless	SMYEV	This study
53 / KP284159	Argentina	F. x ananassa cv. Camarosa	Single/symptomless	SMYEV	This study
264 / KP284160	Argentina	F. x ananassa cv. Albión	Single/symptomless	SMYEV	This study
20 / KP284161	Argentina	F. x ananassa cv. Camarosa	Single/symptomless	SMYEV	This study
Berra2 / in process	Argentina	F. x ananassa cv. Camarosa	Mixed/foliar symptoms	SMYEV, SMoV, SCV	Unpublished
1CH / AJ577337	Chile	*F. chiloensis*	Not available	SMYEV*	Thompson and Jelkmann 2004
2CH / AJ577338	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
3CH / AJ577339	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
4CH / AJ577340	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
5CH / AJ577341	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
6CH / AJ577342	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
7CH / AJ577343	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
D/L.9 / AJ577344	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
9Redland / AJ577345	Australia	F. x ananassa cv. Redlands crimson	Not available	SMYEV*	Thompson and Jelkmann 2004
10CH / AJ577346	Chile	F. chiloensis	Not available	SMYEV*	Thompson and Jelkmann 2004
D/L.13 / AJ577347	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
D/L.14 / AJ577348	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
D/L.19 / AJ577349	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
69 N / AJ577350	Belgium	F. x ananassa cv. unknown	Not available	SMYEV*	Thompson and Jelkmann 2004
314CP2cav / AJ577351	Italy	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
D/M.110 / AJ577352	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
D/K.159 / AJ577353	Germany	F. x ananassa cv. EM834	Not available	SMYEV*	Thompson and Jelkmann 2004
D/V.180 / AJ577354	Germany	F. x ananassa cv. Elsanta	Not available	SMYEV*	Thompson and Jelkmann 2004
1182–53F / AJ577355	United States	F. x ananassa selection 1187	Not available	SMYEV*	Thompson and Jelkmann 2004

Table 1 (continued)

Name of isolate / accession number	Country	Host	Infection/symptoms	Virus detected	Reference
IndukaA / AJ577356	The Czech Republic	*F. vesca* clone UC5	Not available	SMYEV*	Thompson and Jelkmann 2004
IndukaB / AJ577357	The Czech Republic	F. vesca clone UC5	Not available	SMYEV*	Thompson and Jelkmann 2004
WSU1988 / AJ577358	United States	F. x ananassa selection WSU1988	Not available	SMYEV*	Thompson and Jelkmann 2004
D74 / AJ577359	Germany	F. vesca clone UC5	Not available	SMYEV*	Thompson and Jelkmann 2004
MY-18 / D12515	United States	F. x ananassa cv. Totem	Not available	SMYEV*	Thompson and Jelkmann 2004
MY-18 / NC003794	United States	F. x ananassa cv. Totem	Not available	SMYEV*	Thompson and Jelkmann 2004
KNS1 / EU284709	Korea	F. x ananassa cv. Seolhyang	Not available	SMYEV*	Cho et al. 2011
sy01 / AY955375	China	not available	Not available	SMYEV*	Yang and Zhang unpublished
sy02 / EU107084	China	not available	Not available	SMYEV*	Yang and Zhang unpublished
sy03 / EU107085	China	not available	Not available	SMYEV*	Yang and Zhang unpublished
sy04 / EU107086	China	not available	Not available	SMYEV*	Yang and Zhang unpublished

amplification conditions for the nested PCR were the same as those of the first PCR.

Sequence analyses

Amplified PCR products were cloned using the pGEM-*Easy* vector Systems I cloning Kit according to the manufacturer's instructions (Promega Madison, WI, USA). Standard procedures, using blue/white selection, were followed to screen and select transformants (Sambrook et al. 1989). Recombinant plasmids were assayed for inserts by direct PCR amplification of the fragments, as explained above, from bacterial colonies with specific primers YT1/Y2 (Thompson et al. 2003; Thompson and Jelkmann 2004). Positive transformants were grown overnight in LB medium with ampicillin and the plasmids were purified using Plasmid mini Kit (Qiagen, Germany). At least three clones of each SMYEV PCR product of each isolate were completely sequenced on both strands using an ABI 3130xl (Applied Biosystems).

The nucleotide sequences of the gene enconding the CP of the different SMYEV clones were analyzed with Lasergene 8.0.2 package (DNASTAR Inc., Madison, WI, USA). Sequences were obtained and then translated into a protein using EditSeq and SeqMan software of the same package. A total of 12 consensus sequences of clones with over 99.5 % identity within each isolate were generated and used for analysis (Table 1). The obtained nucleotide sequences and deduced amino acids of the CP of Argentinian isolates were compared with 30 SMYEV sequences published in GenBank (Table 1). Forty-two sequences were aligned using CLUSTAL W (Thompson et al. 1994). Identity percentages of nucleotide and amino acid sequences were calculated using MegAlign (DNASTAR Inc., Madison, WI, USA). Phylogenetic trees for the CP gene of all 42 SMYEV isolates were constructed using the Neighbor-Joining method using Mega 6 (Tamura et al. 2013). The CP gene of a PVX isolate (AF485891) was used as out-group. Statistical significance of the tree was estimated by applying a bootstrap of 1000 replicates.

Based on the obtained phylogenetic tree, the genetic variability of the SMYEV subgroups was analyzed to obtain data on nucleotide diversity within and among subgroups (mean number of nucleotide differences per site between two sequences), number of polymorphic sites and, number of mutations, single variants, parsimony informative sites, number of monomorphic sites, number of synonymous and nonsynonymous substitutions. These analyses were performed using D*na*SP v5 software (Librado and Rozas 2009).

Recombination analysis

An analysis was performed to detect the probable recombinant sequences, to identify possible parental sequences and to locate possible recombination spots based on the multiple alignment of CP nucleotide sequence of the Argentinian isolates and those deposited at the GenBank, using the RDP3 software (Martin et al. 2010). This program has nine algorithms for the detection of recombinants: RDP, GENECONV, MAX CHI, CHIMAERA, BOOT SCAN, 3SEQ, SISCAM, PhylPro and LARD. The analysis was performed considering a highest acceptable probability P - value =0.001; the linear sequence option was selected and the remaining parameters were used

with default settings. Predicted recombination breakpoints were accepted when an event was supported by at least three different methods.

Results

Molecular and serological analyses revealed the presence of SMYEV in symptomless strawberry plants. The presence of SMYEV together with SMoV and/or SCV was detected in symptomatic plants, confirming the occurrence of mixed infections. Other viruses were not detected in the analyzed samples.

Sequence analysis

The complete CP sequences of 12 Argentinian SMYEV isolates were determined and deposited in GenBank with accession numbers KP284152 to KP284162. In strawberry plants 13, 16 and 36 two and three isolates were obtained from one plant (13–2-5; 13–3-4; 16–1; 16–2; 16–4-5; 36–1-3; 36–2–4-6). The CP gene of Argentinian isolates was 729 nt in length and encoded a protein of 242 aa, excluding the stop codon. Nucleotide identity and amino acid identity of the CP gene was 81.5–99.6 % and 91.4–100 % among Argentinian isolates, respectively, and 81.6–99 % and 86–100 % with the 30 SMYEV sequences recorded in GenBank, respectively (Table 1).

Phylogenetic analysis

The phylogenetic analysis of the 42 CP sequences allowed the identification of two SMYEV groups (A and B). The first group (A) was divided in two subgroups with bootstrap value 88 % (I and II), and the second group (B) was divided in two subgroups (III and IV) with lower bootstrap value 76 % (Fig. 1). All the Argentinian isolates were distributed across the four subgroups. Subgroup I comprised the highest number of SMYEV isolates (23:12 European, 5 Argentinian, 3 Asian, 2 USA and 1 Chilean). Identity percentages within subgroup I ranged from 95.9 to 100 % of nt and from 96.3 to 100 % of aa. Subgroup II was composed of SMYEV 8 isolates (4 Chilean, 3 Argentinian and 1 Australian) with 94.9–98.9 % nt identity and 96.7–99.2 % aa identity. Subgroup III included 6 SMYEV isolates (3 Chilean, 2 USA and 1 Argentinian) with 97.9–100 % nt identity and 97.5–100 % aa identity among them. Subgroup IV comprised 5 SMYEV isolates (3 Argentinian and 2 Chinese) with 91.1–99.3 % nt identity and 89.7–99.2 % aa identity. Plants 13, 16 and 36 were infected by more than one viral isolate each, which belonged to different subgroups of the phylogenetic tree. The isolates 13–2, 16–1 and 36–1-3 were included in subgroup IV; 13–3-4 and 36–2–4-6 were placed in subgroup I, whereas 16–2 and 16–4-5, which were obtained from the same sample, were included in subgroup II.

No insertions or deletions were observed in the CP genes of the Argentinian isolates; however, nucleotide substitutions were detected in the entire genome segment analyzed. A total of 212 aa out of 242, were identical, and 30 were variable: 11 were conserved substitutions, 4 were semi-conserved substitutions and 15 were different. The start codon ATG, which codes for the CP gene, was observed in all the isolates. The stop codon TAA was detected in all the isolates, except for an isolate from Australia (CP-AJ577345), whose stop codon was TAG. Another difference was observed with respect to the Chinese isolate (CP-EU107085), which has a mutation that results in a nonfunctional CP. Multiple alignment of the aa sequences of the CP gene of all isolates from Argentina as well as from the rest of the world revealed the presence of the conserved sequence FAAFDFFDGV, located in the C-terminal region of the CP of SMYEV.

Multiple alignment of aa of all sequences showed 61 variable aa, of which 16 were conserved substitutions, 10 were semi-conserved substitutions and 35 were different; in the latter case, substitutions were distributed along the virus CP, but most of them were located on the N-terminal region.

Diversity within each subgroup (I, II, III and IV) ranged between 0.0185 and 0.0609, with a maximum standard deviation of 0.01; a high diversity value (0.0609) was recorded only in subgroup IV. The analysis considering all sequences together showed a nucleotide diversity of 0.0923. Sequence diversity among subgroups was estimated as a mean number of nucleotide substitutions between sequence pairs, considering the effect of polymorphism, resulting in values ranging from 0.0896 to 0.1781 (Table 2).

The analyzed descriptors of genetic variability showed differences that varied with the data included in the analyses (Table 3). The analysis including all the descriptors showed 32.5 % of polymorphic sites and 27 % of parsimony-informative sites. For each subgroup, the polymorphic and parsimony-informative sites were I: 12.2 % and 4.8 %, II: 8.9 and 2.05 %, III: 4.2 % and 1.1 % and IV: 11.1 % and 7.7 %, respectively. Synonymous substitutions in the sequences were higher than non-synonymous ones in all cases analyzed.

Recombination analysis

The results of the recombination analysis based on 42 sequences of CP of SMYEV revealed evidences of two potential recombination events supported with the *P* value, based on the program algorithms (Table 4). The recombination event 1 showed recombination between the Argentinian isolate CP-20 (KP284161), as the major parent (97.8 % subgroup II)

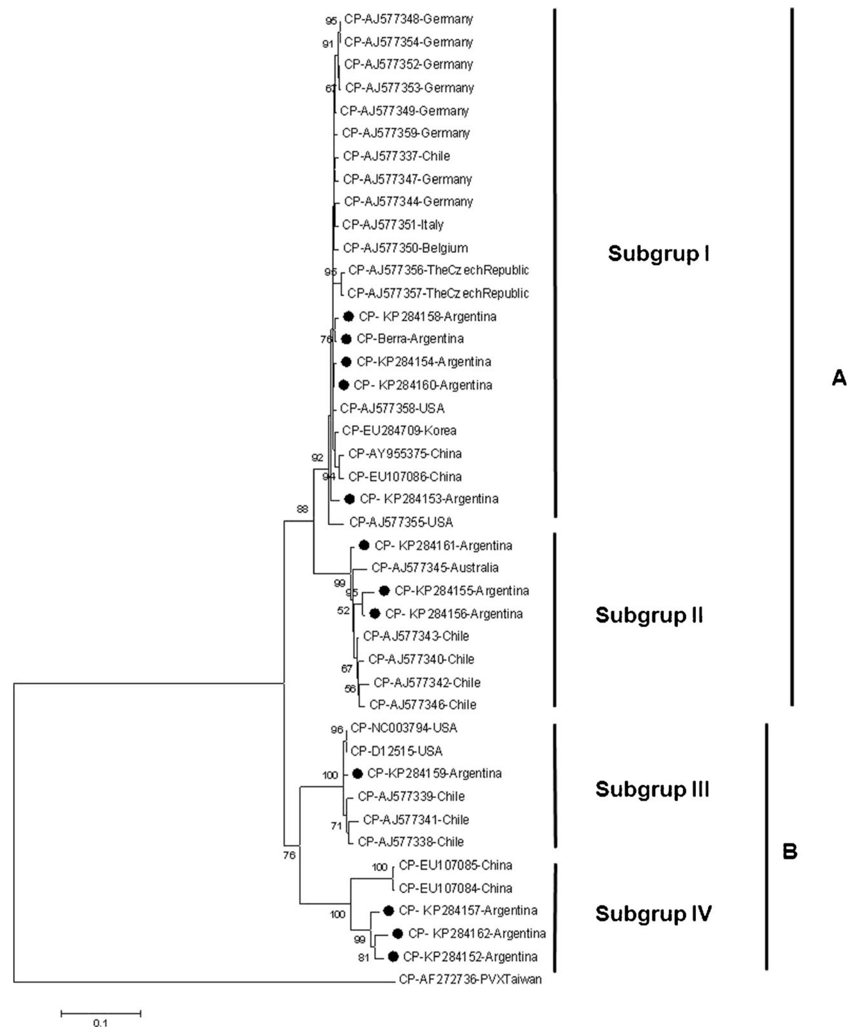

Fig. 1 Neighbor-Joining tree based on the nucleotide sequence alignment of coat protein of different SMYEV isolates, depicting phylogenetic relationships. The number of bootstrap replicates used was 1000. Branches with bootstrap values >50 % are shown. The *scale bar* represents the genetic distance (substitutions per nucleotide)

and the Argentinian isolate CP-16-1 (KP284162) as the minor parent (100 % subgroup IV), which led to a recombinant isolate of Argentina CP-16-2 (KP284155), this recombination event, comprising a region from 1 to 692 nt of the major parent and a region from 693 to 729 nt of the minor parent. The recombination event 2 showed recombination between the Chilean isolate (CP-AJ577346) as the major parent (98.9 % identity, subgroup II) and an unknown minor parent sequence similar to CP-NC003794 (this parent was deduced, but not detected, by the software RDP3), which led to a recombinant isolate of Chilean (CP-AJ577342). In this recombination event, a region from 16 to 82 nt of major parent was replaced

Table 2 Nucleotide diversity of the gene encoding the CP within and between sequence subgroups of SMYEV isolates from different regions of the world

Subgroups	I	II	III	IV	All data
I	**0.0185**	0.0896	0.1372	0.1705	-
II	-	**0.0266**	0.1365	0.1781	-
III	-	**-**	**0.0167**	0.1529	-
IV	-	-	-	**0.0609**	-
All data	-	-	-	-	**0.0923**

Values in bold correspond to within-subgroup diversity

Table 3 Statistical descriptors of genetic variability of SMYEV isolate subgroups

Descriptors	I	II	III	IV
N° of sequences	23	8	6	5
N° of polymorphic sites	89	65	31	81
N° of mutations	89	65	31	83
Single variants	53	50	23	23
N° parsimony informative sites	35	15	8	56
N° of monomorphic sites	639	664	698	646
Synonymous substitutions	65	43	24	66
Nonsynonymous Substitutions	21	21	7	16

Table 4 Recombination events detected in the coat protein gene of the 42 SMYEV isolates

Recombinant isolate	Parents		Breakpoints		Detection methods	P - value
	Major/% identity	Minor/% identity	Start	End		
CP-16-2 KP284155 Argentina	CP-20 KP284161 Argentina/97.8	16–1 KP284162 Argentina/100	692	729	RDP	3.85×10^{-06}
					GENECONV	3.60×10^{-06}
					BOOTSCAM	3.06×10^{-07}
					MAX CHI	1.26×10^{-01}
CP-AJ577342 Chile	CP-AJ577346 Chile/98.9	(NC003794) unknown/no data	16	82	RDP	2.63×10^{-05}
					GENECONV	2.29×10^{-04}
					BOOTSCAM	1.028×10^{-03}
					3Seq	7.71×10^{-04}
					MAX CHI	2.98×10^{-01}

with the corresponding genome sequence of unknown isolate, minor parent (similar to CP-NC003794).

Discussion

The SMYEV nucleotide sequence variability was confirmed since four phylogenetic groups were detected comparing 42 virus isolates from different regions of the world.

In 2004, Thompson and Jelkmann (2004) conducted a phylogenetic analysis of the nucleotide sequences of the CP genes of different SMYEV isolates and found three subgroups; however, the phylogenetic analysis of all the sequences of the CP gene published in GenBank and the Argentinian isolates from this study, suggest four subgroups (I, II, III and IV). Support for the four subgroups is given by the analysis of diversity, which indicates a high degree of conservation within the subgroups, especially I, II and III, but a higher level of diversity among subgroups.

The diversity analysis indicated that subgroup IV had greater variability among the component sequences compared to other groups, which was also evidenced by the greater number of synonymous changes. This is evident in the separation of the Argentinian isolates from the Chinese isolates, with 100 % significance and the lowest within subgroup nucleotide identity values suggesting a probably fifth subgroup. Considering the differences between Argentinian and Chinese strawberry cultivars, these differences could be due to adaptation of the virus to new environmental conditions. RNA viruses are known to have a high potential for genetic variation due to the high mutation and recombination rates, which increases their adaptation ability to new environments. However, the high genetic conservation is a general rule for most of the RNA virus populations infecting plants (García Arenal et al. 2001).

All the commercial strawberry germplasm cultivated in Argentina is imported from nurseries located in other countries, mainly the USA (California) and Europe. Importation of infected material from a range of locations may provide another explanation for the presence of Argentinian SMYEV variants occurring in each of the four subgroups. Germplasm is then vegetatively propagated in nurseries located in the provinces of Chubut, Mendoza and Neuquén and further distributed to growers in different provinces, including Tucumán (Lules), where the isolates compared in this study came from. The presence of the vectors *Chaetosiphon fragaefolii*, *C. thomasi* and *C. minor*, which were detected in different growing regions of Argentina, certainly plays an important role in virus distribution (Ortego 1997; Delfino et al. 2007; Cédola and Greco 2010) and might be responsible for the dispersal of the virus variants in all subgroups.

In addition, in three Argentinian strawberry plants used in this work (13, 16, and 36), two and three isolates of the SMYEV were found in the same plant. These viruses variant were found in both single SMYEV infection and mixed infections with others viruses. Furthermore, the sequences obtained from one plant were included in different subgroups of the phylogenetic tree, suggesting the presence of mixed infections of SMYEV variants in the same plant. A similar result was reported for other viruses, as *Papaya ringspot virus* (*Potyvirus*), and *Wheat streak mosaic virus* (Tritimovirus), with infections of up to three variants in a single plant sample (Abdalla and Ali 2012; Robinson and Murray 2013).

The constant exchange of strawberry cultivars between countries could explain the lack of a distribution of viral variants by geographical regions, as it was explained for the Argentina case. Despite this, all isolations from Europe were detected in the subgroup I, subgroup II and III included principally isolates from Chile and Argentina, and the subgroup IV included isolates from China and Argentina, perhaps supporting the hypothesis of a possible adaptation to local environmental conditions.

The intraspecific recombination analysis of 42 isolates from the SMYEV sequences detected two recombinants, CP16–2 (KP284155) and a Chilean variant (AJ577342). Both recombinants occur in subgroup II because the major parental variant, CP20 (KP284161) and Chilean variant (AJ577346), respectively, also occur subgroup II. However, CP16–2 (KP284155) and Chilean variant (AJ577342) carry short segments of the CP gene from variants in subgroup IV (36 nt, CP16–1) and subgroup III (66 nt, NC003794), suggesting the existence of intraspecific diversity in SMYEV isolates and confirming an exchange of SMYEV genetic material among infected plants, probably induced by vector activity.

The presence of intraspecific recombination events has also been reported in other *Potexvirus* species including the polymerase gene region, the triple gene block and the C-terminal region of the CP gene (Sherpa et al. 2007; Hasiów Jaroszewska et al. 2010). The recombination region in both recombinants found in this study also occurred at the 3′ and 5` terminal CP regions. The central region of the CP gene is more conserved than these regions, which is likely to be due to the conservation of functionally important residues in this part of the encoded protein (Rogers et al. 2012).

Although the RDP3 program predicted two SMYEV recombinants out of 42 isolates, based on analysis of the CP gene, it is possible that an analysis of a larger portion of the genome could have identified further isolates in which recombination events occurred elsewhere in the genome. Conversely, it may be that recombination occurs frequently but most are eliminated from viral populations because they are poorly adapted to the environment (Jaag and Nagy 2010).

The evolutionary process of intraspecific recombination is common in single-stranded positive-sense RNA plant viruses (Chare and Holmes 2006). It is an evolutionary mechanism that provides an additional mean to increase genetic variability in viruses, but it is also a method that viruses with high mutation rates use for "repairing" deleterious mutations (García Arenal et al. 2001; Jaag and Nagy 2010). Recombination has been associated with the increase of viral pathogenicity, host range and capacity to overcome resistance in crop varieties (García Arenal and McDonald 2003). Further studies are needed to evaluate the influence of recombination and elucidate the underlying mechanism in SMYEV.

Acknowledgments This study was carried out at IPAVE-CIAP-INTA and was supported by INTA and CONICET and Fondo para la Investigación Científica y Tecnológica (FONCyT), "Agencia Nacional de Promoción Científica y Tecnológica" grand. The authors are grateful to Drs. Sergio Salazar and Natalia Meneguzzi from EEA Famaillá, INTA, for their valuable support in plant collection.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

References

Abdalla O, Ali A (2012) Genetic diversity in the 3′-terminal region of papaya ringspot virus (PRSV-W) isolates from watermelon in Oklahoma. Arch Virol 157:405–412. doi:10.1007/s00705-011-1184-8

Cédola C, Greco N (2010) Presence of the aphid, Chaetosiphon fragaefolii, on strawberry in Argentina. J Insect Sci 10:1–9. doi:10.1673/031.010.0901

Chang L, Zhang Z, Yang H, Li H, Dai H (2007) Detection of strawberry RNA and DNA viruses by RT-PCR using total nucleic acid as a template. J Phytopathol 155:431–436. doi:10.1111/j.1439-0434.2007.01254.x

Chare ER, Holmes EC (2006) A phylogenetic survey of recombination frequency in plant RNA viruses. Arch Virol 151:933–946. doi:10.1007/s00705-005-0675-x

Cho JD, Kim JS, Lee SH, Chung BN, Choi GS (2011) Strawberry mild yellow edge potexvirus from strawberry in Korea. Plant Pathol J 27: 187–190. doi:10.5423/PPJ.2011.27.2.187

Conci VC, Torrico AK, Cafrune EE, Quevedo V, Baino O, Ramallo JC, Borquez AM, Mollinedo VA, Agüero JJ, Kirschbaum DS (2009) First report of strawberry mild yellow edge virus in Argentina. Acta Hortic 842:303–306

Delfino M, Conci VC, Dughetti AC (2007) Áfidos transmisores de virus de frutilla en la Argentina in: XXX Congreso Argentino de Horticultura. La Plata, Buenos Aires, Argentina p 106

García Arenal F, McDonald BA (2003) An analysis of the durability of resistance to plant viruses. Phytopathology 93:941–952

García Arenal F, Fraile A, Malpica JM (2001) Variability and genetic structure of plant virus populations. Annu Rev Phytopathol 39: 157–186

García Arenal F, Malpica JM, Fraile A (2008) Análisis de la variabilidad de los virus de plantas. In: Pallas V, Escobar C, Rodriguez Palenzuela P, Marcos JF (eds) Herramientas Biotecnológicas En Fitopatología, Madrid, pp 135–146

Hasiów Jaroszewska B, Kuzniar A, Peters S, Leunissen J, Pospieszny H (2010) Evidence for RNA recombination between distinct isolates of Pepino mosaic virus. Acta Biochim Pol 57:385–388

Jaag HM, Nagy PD (2010) The combined effect of environmental and host factors on the emergence of viral RNA recombinants. PLoS Pathog 6:e1001156. doi:10.1371/journal.ppat.1001156

King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. Elsevier Academic Press, San Diego

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452. doi:10.1093/bioinformatics/btp187

Maas JL (1998) Compendium of strawberry diseases. APS Press, USA

Martin D, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462–2463. doi:10.1093/bioinformatics/btq467

Ortego J (1997) Pulgones de la Patagonia Argentina con la descripción de Aphis Intrusa sp. n. (Homoptera: Aphididae). Rev la Fac Agron 102: 59–80

Posthuma KI, Adams AN, Hong Y, Kirby MJ (2002) Detection of strawberry crinkle virus in plants and aphids by RT-PCR using conserved L gene sequences. Plant Pathol 51:266–274. doi:10.1046/j.1365-3059.2002.00725.x

Robinson MD, Murray TD (2013) Genetic variation of wheat streak mosaic virus in the United States Pacific northwest. Phytopathology 103:98–104. doi:10.1094/PHYTO-05-12-0108-R

Rogers SM, Payton M, Allen RW, Melcher U, Carver J, Fletcher J (2012) Method: a single nucleotide polymorphism genotyping method for wheat streak mosaic virus. Investig Genet 3:1–11. doi:10.1186/2041-2223-3-10

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York

Sherpa AR, Hallan V, Pathak P, Zaidi AA (2007) Complete nucleotide sequence analysis of cymbidium mosaic virus Indian isolate: further evidence for natural recombination among potexviruses. J Biosci 32:663–669

Sui C, Wu L, Zhang Z (2003) Detection of strawberry vein banding virus in strawberries by PCR. Acta Hort Sin 820:82

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. doi:10.1093/molbev/mst197

Thompson JR, Jelkmann W (2003) The detection and variation of strawberry mottle virus. Plant Dis 87:385–391

Thompson JR, Jelkmann W (2004) Strain diversity and conserved genome elements in strawberry mild yellow edge virus. Arch Virol 149:1897–1909. doi:10.1007/s00705-004-0353-4

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Thompson JR, Wetzel S, Klerks MM, Vašková D, Schoen CD, Špak J, Jelkmann W (2003) Multiplex RT-PCR detection of four aphid-borne strawberry viruses in Fragaria spp. in combination with a plant mRNA specific internal control. J Virol Methods 111:85–93. doi:10.1016/S0166-0934(03)00164-2

Tzanetakis IE, Martin RR (2013) Expanding field of strawberry viruses which are important in North America. Int J Fruit Sci 13:184–195. doi:10.1080/15538362.2012.698164