



**HAL**  
open science

## **New species in the papaya ringspot virus cluster: insights into the evolution of the PRSV lineage**

Cecile Desbiez, Catherine Wipf-Scheibel, Pauline Millot, Eric Verdin, Gasim  
Dafalla, Hervé Lecoq

### ► **To cite this version:**

Cecile Desbiez, Catherine Wipf-Scheibel, Pauline Millot, Eric Verdin, Gasim Dafalla, et al.. New species in the papaya ringspot virus cluster: insights into the evolution of the PRSV lineage. *Virus Research*, 2017, 241, pp.88-94. 10.1016/j.virusres.2017.06.022 . hal-01602711

**HAL Id: hal-01602711**

**<https://hal.science/hal-01602711>**

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

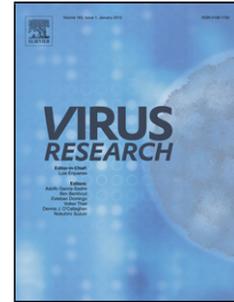


Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

## Accepted Manuscript

Title: New species in the papaya ringspot virus cluster:  
insights into the evolution of the PRSV lineage

Authors: C. Desbiez, C. Wipf-Scheibel, P. Millot, E. Verdin,  
G. Dafalla, H. Lecoq



PII: S0168-1702(17)30164-8  
DOI: <http://dx.doi.org/doi:10.1016/j.virusres.2017.06.022>  
Reference: VIRUS 97178

To appear in: *Virus Research*

Received date: 17-2-2017  
Revised date: 24-5-2017  
Accepted date: 26-6-2017

Please cite this article as: Desbiez, C., Wipf-Scheibel, C., Millot, P., Verdin, E., Dafalla, G., Lecoq, H., New species in the papaya ringspot virus cluster: insights into the evolution of the PRSV lineage. *Virus Research* <http://dx.doi.org/10.1016/j.virusres.2017.06.022>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

New species in the papaya ringspot virus cluster: insights into the evolution of the PRSV lineage

Desbiez C<sup>1</sup>, Wipf-Scheibel C<sup>1</sup>, Millot P<sup>1</sup>, Verdin E<sup>1</sup>, Dafalla G<sup>2</sup>, Lecoq H<sup>1</sup>

<sup>1</sup> INRA, UR407, Unité de Pathologie Végétale, 84140 Montfavet, France

<sup>2</sup> Plant Pathology Center, University of Gezira, Wad Medani, Sudan

Highlights :

- Two new potyvirus species from Sudan were defined in the PRSV cluster
- SuWMV is an interspecific recombinant in the 5' part of its genome
- The PRSV cluster may have radiated about 3600 years ago
- The PRSV cluster underwent an important diversification in Africa

Abstract

The “Papaya ringspot virus (PRSV) cluster” of cucurbit-infecting potyviruses contains five acknowledged species that have similar biological, serological and molecular properties. Additional data suggest there are other uncharacterized species from various locations in the world that likely belong to the PRSV cluster including a new PRSV-like virus reported from Sudan in 2003. Molecular and biological data indicated that the virus from Sudan belongs to a new species, tentatively named wild melon vein banding virus (WMVBV). The complete nucleotide sequence of a second virus from Sudan revealed it was a divergent relative of *Moroccan watermelon mosaic virus* (MWMV). Based on sequence similarity this virus was determined to be a distinct species and tentatively named *Sudan watermelon mosaic virus*

(SuWMV). Molecular analyses indicate that SuWMV is a recombinant between WMVBV- and MWMV-related viruses. Based on surveys performed in Sudan between 1992 and 2012, SuWMV appeared 10 times more frequent than WMVBV in that country (14.6% vs. 1.5% of the samples tested). The geographic structure and molecular diversity patterns of the putative and acknowledged species suggest that the PRSV-like cluster originated in the Old World about 3600 years ago, with an important diversification in Africa.

Keywords: potyvirus, interspecific recombination, cucurbit, evolution

## 1. Introduction

The genus *Potyvirus* contains about 25% of all known plant RNA viruses. Molecular criteria are the primary means to distinguish species with nucleotide sequence identity of the coat protein or the whole genome of less than 76% defining a new species (Adams et al., 2005b). Potyvirus star-like phylogeny suggests an initial radiation in the genus that, based on a molecular clock calibration using historical isolation and outbreak events, took place about 6600 years ago, i.e at the dawn of agriculture (Gibbs et al., 2008a). Several “clusters” of closely related species have been defined within the Potyvirus genus with viruses in a cluster sharing 70 to 80% sequence identity in the coat protein coding region (Gibbs et al., 2008b; Romay et al., 2014; Spetz et al., 2003). Therefore, the demarcation between the different species is not always clear-cut. Viruses within a cluster usually have overlapping host ranges and are, in some cases, affected by the same resistance genes.

These features characterize the viruses in the “papaya ringspot virus (PRSV) cluster” (Quiot-Douine et al., 1990; Romay et al., 2014). Using biological and serological data on 20 PRSV-related isolates, Quiot-Douine et al. defined at least seven related isolate groups (Quiot-Douine et al., 1990). Subsequent molecular analysis indicated that five of these seven groups correspond to acknowledged species: PRSV, *Moroccan watermelon mosaic virus* (MWMV) (Lecoq et al., 2001), *Algerian watermelon mosaic virus* (AWMV) (Yakoubi et al., 2008b), *Zucchini yellow fleck virus* (ZYFV) (Desbiez et al., 2007) and *Zucchini tigre mosaic virus* (ZTMV) (Romay et al., 2014). Other species, different from the seven previously studied groups, have subsequently been characterized in the PRSV cluster. *Zucchini shoestring virus* (ZSTV), was recently described as a new species from South Africa (Ibaba et al., 2015) and we report here on additional PRSV-like virus isolates from Sudan. Furthermore, based on partial sequence, the taxonomic status of a highly divergent isolate of MWMV from Sudan remains unresolved (Lecoq et al., 2001; Adams et al., 2005b). In this study, we assessed the taxonomic position of two PRSV-like isolates from Sudan, Su94-54 and Su03-07, characterized the prevalence and molecular diversity of PRSV-related viruses in Sudan and analyzed the genetic variability and geographic distribution of the PRSV-related viruses in order to understand better the evolution and spread of the PRSV lineage.

## 2. Materials and methods

### 2.1 Virus isolates

Su94-54 was collected in 1994 in Eastern Sudan near Gadamballia from a snake cucumber (*Cucumis melo* var. *flexuosus*) showing symptoms of mosaic and leaf deformations and was initially described as a highly divergent isolate of MWMV (Lecoq et al., 2001). Dried leaf material infected with the virus was stored at 4<sup>o</sup> C over calcium chloride. To preserve the

infectivity over time, extracts from the dry material were ground in phosphate solution and mechanically inoculated on zucchini squash as described by Lecoq et al., 2001, at variable intervals since 1994. Material from the newly infected plants was dried and stored at 4° C.

Su03-07 was collected in Tayba Block (Gezira Scheme, South-West of Wad Medani between the Blue Nile and White Nile Rivers), Sudan during a survey of cucurbit species in December 2003. The wild melon *Cucumis melo* var. *agrestis* (Su03-07) exhibited mild vein banding and mosaic symptoms, but DAS-ELISA using panel of 10 antisera to cucurbit-infecting viruses (Lecoq et al., 2016) were all negative. When mechanically inoculated to zucchini squash, the inoculated plants developed systemic chlorotic spots within two weeks post inoculation. The original isolate, as well as tissue from the inoculated zucchini squash plants, were stored as dried material on calcium chloride.

## 2.2 Host range of Su03-07

Two plants of zucchini squash were mechanically inoculated with dried material of isolate Su03-07. One positive plant was used as source for mechanical inoculation to young plantlets of several cucurbit and non-cucurbit species: *Cucumis melo* cultivars Védreantais, Ouzbèque and WMR29, *C. sativus* (cucumber) cv. Beit Alpha, *Cucurbita pepo* (zucchini squash) cv. F1 Diamant, *Citrullus lanatus* (watermelon) cv. Sugar Baby, *Chenopodium amaranticolor*, *C. quinoa*, *Nicotiana benthamiana*, *N. tabacum* cv. Xanthi, *Solanum lycopersicum* (tomato) cv. Monalbo, *Capsicum annuum* (pepper) cv. Yolo Wonder, petunia x hybrida, *Phaseolus vulgaris* (bean) cv. Pinto and *Pisum sativum* (pea) cv. Douce Provence. The plants were kept for 4 weeks in insect-proof greenhouse. All asymptomatic plants were tested by RT-PCR using specific primers for Su03-07 (see below). For symptomatic hosts, only one plant of each species or cultivar was tested.

### 2.3 Complete sequencing of Su03-07 and Su94-54

Total RNA from isolate Su03-07 was extracted from infected zucchini squash using TRI-reagent (Molecular Research Center, Cincinnati, OH) following the manufacturer's recommendations. Reverse transcription, PCR amplification and sequencing of Su03-07 genome were performed as previously described for members of the PRSV clusters, using 3 pairs of "potyvirus" primers (Colinet et al., 1998; Gibbs and Mackenzie, 1997; Ha et al., 2008), one primer PRSV-5NC-5' designed at the 5' extremity of several members of the PRSV cluster (Romay et al., 2014; Yakoubi et al., 2008a; Yakoubi et al., 2008b), and specific internal primers as needed, designed from the partial sequences obtained with the generic primers. Sequence assembly was performed with CAP3 (Huang and Madan, 1999) with manual corrections when needed.

To obtain the complete sequence of isolate Su94-54, total RNA was extracted from dried leaf material using PureLink® Plant RNA Reagent (Ambion™, Life Technologies), and sent for HiSeq sequencing to Fasteris SA, Switzerland for small RNA library construction and deep sequencing after acrylamide gel size purification of 18-30 nt small RNAs. Fragments were assembled with CLC Workbench 7 (CLC Bio, Aarhus, Denmark). The reads were then aligned to the obtained contig, in order to estimate the depth of coverage.

### 2.4 Sequence alignments

Sequence comparisons were performed with BLASTn (Altschul et al., 1997). The taxonomic position of Su03-07 and Su94-54, based on full-length nucleotide sequence comparison with 942 members of the *Potyviridae* family, was estimated with PASC (<http://www.ncbi.nlm.nih.gov/sutils/pasc/>). The complete nt sequences of Su03-07 and Su94-54 were aligned against the sequences of 24 PRSV-related viruses, including at least one

member of each acknowledged species for which complete sequences are available (Fig. 1): four MWMV isolates (EF579995, KU315175-176, LN810061), four ZTMV isolates (KC345605-608), the recently described ZSTV from South Africa (Ibaba et al., 2015, KU355553), AWMV (EU410442) and 14 PRSV isolates (AY162218, AY027810, DQ340771, X97251, EF183499, AY010722, AY231130, EU126128, X67673, DQ374153, DQ374152, EF017707, EU475877, KC345609) using ClustalW included in MEGA6 (Tamura et al., 2013). For coding regions, nucleotide alignments were performed from the deduced amino-acid alignments.

### *2.5 Prevalence and molecular variability of PRSV-like viruses in Sudan*

Seven hundred and fifty cucurbit samples collected in Sudan between 1992 and 2012, were stored as dried leaf material on calcium chloride, frozen sap extract or total RNAs extracted with TRI-reagent (Molecular Research Center, Cincinnati, OH) (Desbiez et al., 2016; Lecoq et al., 2016). All samples were tested by DAS-ELISA using antisera raised against PRSV, MWMV and the Su94-54 isolate of SuWMV (Lecoq et al., 2001)

Total RNA from ELISA-positive samples was extracted with TRI-reagent. Reverse transcription was performed with primer Su54-CP1-R (5'-GGATTGACTGGAGAAGGTG-3') and PCR was performed with a primer pair Su54-Nib-F (5'-CCGGAGCACAGACTTGAAGC-3') and Su54-CP2-R (5'-CTCGAACTGGCTTACTGTGCGC-3') for SuWMV. Primers PRSV-milCP-5' + PRSV-finCP-3' (Lecoq & Desbiez, 2012) and PRSV-5' + PRSVpoly-3' (Romay et al., 2014) were used for the amplification of PRSV CP.

To detect WMVBV against which no specific antiserum was available, a specific primer pair was designed: Su03-07-Nib-5' (5'-TTYTCATCTCGACACACACG-3') and Su03-07-CP-3' (5'-AANCCATTWAGCACGATTGG-3'), that amplify the N-terminal variable part of the CP. RT-PCR was performed on a subset of 206 out of the 750 Sudanese isolates, including 45, 84 and 29

samples collected in surveys done in 2002, 2003 and 2012, respectively. PCR-positive samples were sent to Genoscreen, France for direct sequencing using primer Su03-07-Nib-5'.

### *2.6 Core Capsid Protein sequence comparisons*

The coherently-evolving region of the CP sequences (Gibbs et al., 2008a), i.e. the core and C-terminal regions minus the repetitive N domain, of 45 isolates in the PRSV cluster were aligned with ClustalW included in MEGA6. The PRSV cluster isolates included members of the five acknowledged species, PRSV, MWMV, AWMV, ZYFV, and ZTMV, as well as SuWMV, ZSTV, WMVBV, and cucurbit-infecting potyvirus isolates from Myanmar, described as divergent PRSV isolates may represent at least two additional uncharacterized species in the PRSV cluster (Romay et al., 2014). Three cucurbit-infecting potyviruses that do not belong to the PRSV cluster, zucchini yellow mosaic virus (ZYMV), watermelon mosaic virus (WMV) and watermelon leaf mottle virus (WLMV) were included in the analysis as outgroups. The Genbank accession number of each isolate is indicated in Fig. 2. Nucleotide alignments were performed from the deduced amino-acid alignments.

### *2.7 Recombination analysis*

Identification of potential genome recombination sites was performed using RDP4 (Martin et al., 2015). This analysis sequentially tests every combination of three sequences in an input alignment for evidence that one of the three sequences is a recombinant and the other two are its parents. RDP4 uses a range of methods including the original RDP method (Martin and Rybicki 2000), BOOTSCAN (Salminen et al. 1995), MAXCHI (Maynard Smith 1992), CHIMAERA (Posada and Crandall 2001), 3SEQ (Boni, Posada, and Feldman 2007), GENECONV (Padidam,

Sawyer, and Fauquet 1999), LARD (Holmes, Worobey, and Rambaut 1999), and SISCAN (Gibbs, Armstrong, and Gibbs 2000). Only recombination signals detected by more than four methods were used to determine approximate breakpoint positions. A first analysis was performed between the complete nt sequences of Su03-07, Su94-54, four MWMV isolates, four ZTMV isolates, ZSTV, AWMV and 14 PRSV isolates. A second analysis focused on Su94-54, Su03-07 and the four complete MWMV sequences. Visual observation of similarities between complete nucleotide sequences of one isolate from each of the seven PRSV-like species or tentative species (PRSV, MWMV, AWMV, ZTMV, ZSTV, Su94-54 and Su03-07) was performed with SimPlot (Ray, 1998) using uncorrected distances, with a window size of 500 nt and a step size of 50 nt. SplitsTree (Huson & Bryant, 2006) was used on the same sequence alignment as above to check for the presence of network phylogenetic relationships between PRSV-like viruses, suggestive of recombination. RDP4 was also used to check for the presence of recombination signals in the coreCP region, selecting only signals detected by at least four different methods.

## 2.8 Phylogenetic and evolutionary analyses

After checking for the absence of recombination signal, the coreCP alignment of 45 PRSV cluster isolates and three outgroups was used for tree reconstruction. Neighbour-joining and maximum likelihood trees were built on aligned sequences using MEGA6.0, with bootstrap resamplings (n=500 bootstraps).

To estimate the time of divergence of the PRSV cluster, eighty-one core CP sequences were aligned. These include 36 virus species and isolates in the PRSV cluster, 11 viruses in the bean common mosaic virus (BCMV) cluster (Gibbs et al., 2008b), 26 potyviruses outside these two

clusters, and eight potyvirids outside the *Potyvirus* genus (Supplementary table S1). Patristic distances were calculated using PATRISTIC v1.0 (Fourment and Gibbs, 2006; Gibbs et al., 2008b), and times of divergence were estimated using a substitution rate of 1.25 nt/site/year (Gibbs et al., 2008a). A Bayesian coalescent framework by a Markov Chain Monte Carlo (MCMC) method was applied to the aligned coreCP dataset using the Bayesian Evolutionary Analysis by Sampling Trees (BEAST v1.8.3) program (<http://beast.bio.ed.ac.uk/>) (Drummond et al., 2012), with an uncorrelated relaxed clock model. Potyviruses were defined as a monophyletic clade with a node calibration at 6600 yrs. The convergence of the parameters to a stationary distribution and the estimation of the time to the most recent common ancestor (TMCRA) for the PRSV and BCMV clusters were assessed with TRACER v1.6.

### 3. Results and Discussion

#### 3.1 Host range of Su03-07

Su03-07 induced a rapid top necrosis on the melon cultivar 'Védrantais', a mosaic and vein banding in cv. 'Ouzbèque', and local lesions and no systemic symptoms in 'WMR29', a PRSV resistant accession (Table 1). These symptoms were similar to those reported for MWMV (Quiot-Douine et al., 1990, Lecoq et al., 2001). Su03-07 induced vein banding in watermelon, cv 'Sugar Baby' and cucumber, cv 'Beit Alpha', and systemic chlorotic spots in Zucchini Squash, cv 'Diamant'. Among non-cucurbit plants, Su03-07 induced chlorotic local lesions without systemic infection on *Chenopodium amaranticolor* and *C. quinoa*. No symptoms were observed and no virus was detected by RT-PCR after inoculation of *Nicotiana benthamiana*, tobacco cv. Xanthi, tomato, pepper, petunia x hybrida, bean and pea (data not shown). Similar to other members of the PRSV-like cluster, Su03-07 appears to be restricted to cucurbits, and

its biological properties are very similar to that of other members of the cluster including PRSV, MWMV and isolate Su94-54 (Table 1).

### 3.2 Complete sequences of Su94-54 and Su03-07

The full-length sequences of isolates Su94-54 and Su03-07 were determined to be 10,232 nt and 10,152 nt, respectively, excluding the poly(A) tails (Genbank accession numbers KY623505 and KY623506, respectively). For Su94-54, 5,824,335 HiSeq reads of 21-24 nt mapped to the contig obtained with CLC, with an average coverage of 12,191. Both Su03-07 and Su94-54 had a typical potyvirus genome organization, with a large ORF encoding a 3281 aa and 3295 aa polyprotein, respectively, with a putative frameshift in the P3 coding region yielding a P3N-PIPO protein (Chung et al., 2008). The putative cleavage sites in the polyprotein were similar to those of other potyviruses (Adams et al., 2005a). The deduced amino-acid sequence of Su03-07 and Su94-54 contained the three conserved motifs involved in aphid transmission of potyviruses (Pirone and Blanc, 1996); “DAG” was present in the N-terminal part of the CP and “PTK” or “PTR” in the HC-Pro protein. Both viruses had a “RITC” in the N-terminal part of the HC-Pro instead of the conserved “KITC”. Su94-54 and Su03-07 were transmissible by the aphid *Myzus persicae* (Lecoq et al., 2001 for Su94-54; data not shown for Su03-07).

### 3.3 Sequence comparisons and recombination analysis on full-length sequences

The full-length Su03-07 nucleotide sequence shared about 60% identity to the full-length sequences of the other members of the PRSV-like cluster, PRSV (61-61.4%), MWMV (61.5-62%), ZTMV (61-61.5%), AWMV (61.2%) and ZSTV (61%) (Supplementary Table S2). Su03-07 sequence identity in core CP ranged from 69 to 73% with the different members of the PRSV-like cluster. There was no evidence for recombination between Su03-07 and other members of the PRSV cluster (data not shown). Given that the threshold for the distinction of potyvirus

species is a nucleotide sequence identity of 76% or greater in the CP or whole genome (Adams et al., 2005), Su03-07 is a new potyvirus species we are tentatively naming Wild melon vein banding virus (WMVBV).

While the core CP nt and aa identities of Su94-54 were previously determined to be 75-77% and 86%, respectively with MWMV (Lecoq et al., 2001), PASC analyses indicated nt identity across the complete genome was around 70%. Su94-54 shared 69.7-70.1% nt identity with MWMV, 63.8% with AWMV, 62.6 with ZSTV, 60.4-60.9% with ZTMV and 60-60.5% with PRSV. Based on the threshold for the distinction of virus species, Su94-54 is a new potyvirus species we are tentatively naming Sudan watermelon mosaic virus (SuWMV)

Across the whole genome, excepting the 5'NC and P1 coding regions, the lowest nt and aa identities between Su94-54 and MWMV were in the P3 (68 and 67% nt and aa identity, respectively) and the highest identities were in the N1b coding region (74 and 88% nt and aa identity, respectively). The P1 coding region of Su94-54 (1587 nt) was much longer than that of MWMV (1035 nt), aligned poorly and had a low percentage of identity with MWMV (46% nt, 40% aa). SplitsTree analysis revealed that Su94-54 appeared related to MWMV and Su03-07 in a network phylogeny (Fig. 1a). Indeed, the 5' untranslated region and beginning of P1 coding region of Su94-54 were more similar in length and sequence to that of Su03-07 than to MWMV (Fig. 1b). The recombination analysis done using the MWMV, Su03-07, and Su94-54 sequences revealed a recombination junction in Su94-54 around position 1080 within the P1 coding region by six of the eight methods used in the RDP4 analysis. There was no evidence for other recombination signals in the dataset.

Pairwise comparisons of the full-length nt sequences of Su94-54 and the 24 viruses examined within the PSRV cluster (see section 2.4) yielded identities of either less than 70.2%

(maximum 70.1% between Su94-54 and MWMV-GR13-DYS5) or more than 83% (Supplementary Table S2). When the P1 coding region was excluded, removing the possible recombination-related bias, nt identities were either below 72% or above 84% (data not shown). Thus, even between virus isolates that share similarities above the threshold level in the CP sequences, and whatever the recombination status of the P1 coding region, full-length or nearly full-length sequences reveal unambiguous distinction of virus species.

### *3.4 Prevalence and molecular variability of PRSV-like viruses in Sudan*

Two hundred and six virus isolates collected in Sudan between 1994 and 2012 were tested for the presence of WMVBV. Two of these isolates tested positive; Su99-18, a snake melon sample collected in 1999 in Abou Frou about 50 km North from Tayba Block where Su03-07 was collected, and Su12-21, a watermelon sample with strong mosaic and vein banding symptoms collected in 2012 in Wad Abutifa, about 150 km South-East from Tayba Block. Partial sequencing (Genbank accessions KY623546 for Su99-18 and KY623547 for Su12-21) revealed that these isolates shared 92.3% and 96.8% nt identity respectively with Su03-07 in the Nib-CP-3NC region. This shows they belong to the same tentative species and that the virus was already present in Sudan at least 18 years ago.

Based on serological tests, the prevalence of SuWMV in the 750 samples collected in Sudan between 1992 and 2012 was around 17% (92/543) between 1992 and 1999, whereas among the 207 samples collected between 2002 and 2012, it was found only in one isolate from 2006. PRSV was detected in 1995 (17 isolates), 1996 (1 isolate) and 2012 (2 isolates), in the Singa and Damazin regions, 250 and 500 km South-East of Khartoum, respectively. Partial sequences were obtained for 45 isolates of SuWMV (see below) and 20 isolates of PRSV (accessions KY623550-558 for the 9 non-identical sequences obtained). The molecular

diversity among PRSV isolates was limited, and they all clustered within the “Asian” clade of PRSV, as already observed for an isolate collected in 2012 in Wad Ramli, about 100 km north of Khartoum (Mohammed et al., 2012) (Fig. 2 and data not shown).

Among the 45 partially sequenced isolates of SuWMV, five were mixed infections with different variants of the virus (data not shown). Among the 40 non-ambiguous sequences (accessions KY623506-545), two well-supported molecular clades were detected (Fig. 3) without obvious relation to the host or the place of sampling. There was a major clade containing isolates from 1993-1996 and 2006, and a minor divergent clade containing only 2 isolates from 1999. The differentiation between the two clades reached 17 % in the 440-446-nt sequenced region encompassing the variable N-terminal part of the CP. There were also two well supported subclades in the major clade, one containing isolates collected in 1993, 1994 and 1996 in the Gash Delta region, the other one containing mostly isolates from 1995 from the Wad Medani area.

### 3.5 Molecular clock analyses

In the coreCP region, patristic distances between the members of the same species reached 0.2 for ZTMV, ZYFV and MWMV, and 0.17 for PRSV as already estimated by Gibbs et al. (Gibbs et al., 2008a). The average divergence between the PRSV cluster and the other potyviruses was 0.9. The substitution rate of potyviruses has been estimated around  $1.25 \cdot 10^{-4}$  substitutions/site/year, with an initial major radiation 6600 yrs ago and an average patristic distance of 1.67 ns/s (Gibbs et al., 2008a). Both estimation of patristic distances and Bayesian analyses using BEAST yielded a divergence time for the PRSV cluster starting about 3600 years ago. Taking into account the limited precision of the molecular clock and divergence estimations, this appears to be in the same order of magnitude as the BCMV cluster (Gibbs et

al., 2008b), probably originating from South-Eastern Asia and the western Pacific. This may also correspond to the time of species differentiation among sobemoviruses (Fargette et al., 2008). This could suggest that a global and so far uncharacterized environmental and/or climatic factor may have affected plant virus evolution at that time.

### 3.6 Geographic distribution and evolution of PRSV-like viruses

Among the members of the PRSV cluster, only PRSV has been reported to occur worldwide. Phylogenetic studies suggest that PRSV originated from India and spread to Asia and the New World about 400 yrs ago (Gibbs et al., 2008a; Olarte-Castillo et al., 2011). ZTMV has been reported from three continents and some Caribbean and Indian Ocean islands (Romay et al., 2014). Isolates from Asia appear basal relative to American and Caribbean isolates (Fig. 2), suggesting that, like PRSV, ZTMV probably originated in the Old World. The other members of the cluster appear more restricted geographically. MWMV is present in Africa (Lecoq and Desbiez, 2012; Yakoubi et al., 2008a) and has been recently introduced in southern Europe (Lecoq and Desbiez, 2012; Malandraki et al., 2015), but it has not been found on other continents (Lecoq and Desbiez, 2012). AWMV was detected once in Algeria in 1986 and once in Nigeria in 2005-2006 (Yakoubi et al., 2008b, Owolabi et al., 2011). ZYFV was observed occasionally in Crete, Israel and Italy in melon crops (Antignus et al., 1995; Avgelis, 1985; Tomassoli et al., 2010) and in France, where it is mostly confined to the wild cucurbits *Ecballium elaterium* and *Bryonia dioica* (Lecoq and Desbiez, 2012).

WMVBV was found only in Sudan and in a rather limited geographic area. Its low prevalence in Sudan contrasts with its maintenance over 13 years, suggesting that this virus is endemic in wild hosts –cucurbits or others, either annual or perennial- and infects cultivated cucurbits only occasionally. Similarly, SuWMV has also only been observed in Sudan. Its

prevalence in surveys performed between 1992 and 1999 was much higher than between 2002 and 2012, although the surveys were performed at similar times of the year and at least some of the sampled areas were common in the different years (Lecoq, unpublished).

Unexpectedly, SuWMV was found to be an intraspecific recombinant between WMVBV-like and MWMV-like viruses. Few interspecific recombinants have been described among potyviruses until now, and except for SuWMV they were only among members of the BCMV cluster (Desbiez & Lecoq, 2004; Larsen et al., 2005; Valli et al., 2007). For all interspecific recombinants described so far, including SuWMV, the genetic exchange concerns the N-terminal part of the P1 coding region. This region shows the highest diversity in length and sequence among potyviruses (Revers & Garcia, 2015). It is one of the least constrained regions of the potyvirus genome (Moury et al., 2006) and is not required for virus viability (Revers & Garcia, 2015). Since it is self-cleaved during proteolytic maturation, it does not require a compatibility of its cleavage site with the rest of the genome. Therefore, recombinants in the N-terminal part of P1 may be more viable than those involving other viral regions. Recombination in the P1 was suggested to facilitate adaptation to a wide range of host species (Valli et al., 2007). Regarding SuWMV, like the other members of the PRSV cluster its host range is mostly restricted to cucurbits, and the symptoms it induces are similar to those of other members of the cluster (Table 1 and Lecoq et al., 2001). Its prevalence in Sudan appears higher than that of its putative parental viruses MWMV and WMVBV, but whether its recombinant nature plays any role in this situation remains unknown.

Despite similar apparent biological and molecular characteristics, members of the PRSV cluster have a very diverse geographic distribution and prevalence. Intrinsic properties of the viruses such as host range (e.g. for PRSV, the ability to infect papaya), seed transmission (Gibbs

et al., 2008a), vectors and virus reservoirs, as well as extrinsic factors such as environment and global trade of plant material, can contribute to the fact that some viruses remain endemic whereas other ones cause important epidemics worldwide. The present emergence of MWMV in southern Europe shows that introduced viruses can become serious agronomic problems, and highlights the importance of characterizing viral pathosystems particularly in the center of origin or diversification of important crops, like Sudan for melon and watermelon, to anticipate future epidemics.

## References

- Adams, M.J., Antoniw, J.F. and Beaudoin, F., 2005a. Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. *Mol. Plant Pathol.* 6, 471-487.
- Adams, M.J., Antoniw, J.F. and Fauquet, C., 2005b. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Arch. Virol.* 150, 459-479.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
- Antignus, Y., Levy, D. and Cohen, S., 1995. Characterisation of a variant of zucchini yellow fleck virus (ZYFV), a potyvirus causing a wilt disease of melons in Israel. *Ann. Appl. Biol.* 126, 111-120.
- Avgelis, A.D., 1985. Epidemiological studies of Zucchini yellow fleck virus in Crete. *Phytopath. Mediterr.* 24, 208-210.
- Boni, M.F, Posada, D. and Feldman, M.W., 2007. An Exact Nonparametric Method for Inferring Mosaic Structure in Sequence Triplets. *Genetics* 176, 1035-1047.
- Chung, B.Y.-W., Miller, W.A., Atkins, J.F. and Firth, A.E., 2008. An overlapping essential gene in the *Potyviridae*. *Proc. Natl Acad. Sci. USA* 105, 5897-5902.
- Colinet, D., Kummert, J. and Lepoivre, P., 1998. The nucleotide sequence and genome organization of the whitefly transmitted sweetpotato mild mottle virus: a close relationship with members of the family *Potyviridae*. *Virus Res.* 53, 187-196.

- Desbiez, C. and Lecoq, H., 2004. The nucleotide sequence of Watermelon mosaic virus (WMV, Potyvirus) reveals interspecific recombination between two related potyviruses in the 5' part of the genome. *Arch. Virol.* 149, 1619-1632.
- Desbiez, C., Justafre, I. and Lecoq, H., 2007. Molecular evidence that zucchini yellow fleck virus is a distinct and variable potyvirus related to papaya ringspot virus and Moroccan watermelon mosaic virus. *Arch. Virol.* 152, 449-455.
- Desbiez, C., Verdin, E., Tepfer, M., Wipf-Scheibel, C., Millot, P., Dafalla, G. and Lecoq, H., 2016. Characterization of a new cucurbit-infecting ipomovirus from Sudan. *Arch. Virol.* 161, 2913-2915.
- Drummond, A.J., Suchard, M.A., Xie, D. and Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969-1973.
- Fargette, D., Pinel-Galzi, A., Sérémé, D., Lacombe, S., Hébrard, E., Traoré, O. and Konaté, G., 2008. Diversification of Rice yellow mottle virus and related viruses spans the history of agriculture from the Neolithic to the Present. *PLoS Pathogens* 4(8), e1000125.
- Fourment, M. and Gibbs, M.J., 2006. PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. *BMC Evol. Biol.* 6(1).
- Gibbs, A. and Mackenzie, A., 1997. A primer pair for amplifying part of the genome of all potyvirids by RT-PCR. *J. Virol. Meth.* 63, 9-16.
- Gibbs, A.J., Ohshima, K., Phillips, M.J. and Gibbs, M.J., 2008a. The prehistory of potyviruses: their initial radiation was during the dawn of agriculture. *PLoS One* 3(6), 1-11.
- Gibbs, A.J., Trueman, J.W.H. and Gibbs, M.J., 2008b. The bean common mosaic virus lineage of potyviruses: where did it arise and when? *Arch. Virol.* 153, 2177-2187.
- Gibbs, M.J., Armstrong, J.S. and Gibbs, A.J., 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16, 573-582.
- Ha, C., Coombs, S., Revill, P.A., Harding, R.M., Vu, M. and Dale, J.L., 2008. Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. *Arch. Virol.* 153, 25-36.
- Holmes, E.C., M Worobey, M. and Rambaut, A., 1999. Phylogenetic evidence for recombination in dengue virus. *Mol. Biol. Evol.* 16, 405-409
- Huang, X. and Madan, A., 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9, 868-877.

- Huson, D.H. and Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254-267.
- Ibaba, J.D., Laing, M.D. and Gubba, A., 2015. First Report of a Novel Potyvirus from the Papaya ringspot virus Cluster Infecting Zucchini (*Cucurbita pepo*) in KwaZulu-Natal, Republic of South Africa. *Plant Dis.* 9, 1289.
- Larsen, R.C., Miklas, P.N., Druffel, K. and Wyatt, S.D., 2005. NL-3 K strain is a stable and naturally occurring interspecific recombinant derived from Bean common mosaic necrosis virus and Bean common mosaic virus. *Phytopathology* 95, 1037-1042.
- Lecoq, H., Dafalla, G., Desbiez, C., Wipf-Scheibel, C., Delecolle, B., Lanina, T., Ullah, Z. and Grumet, R., 2001. Biological and molecular characterization of Moroccan watermelon mosaic virus and a potyvirus isolate from Eastern Sudan. *Plant Disease.* 85, 547-552.
- Lecoq, H. and Desbiez, C., 2012. Viruses of cucurbit crops in the mediterranean region: an ever-changing picture. *Adv. Virus Res.* 84, 67-126.
- Lecoq, H., Verdin, E., Tepfer, M., Wipf-Scheibel, C., Millot, P., Dafalla, G. and Desbiez, C., 2016. Characterization and occurrence of squash chlorotic leaf spot virus, a tentative new torradovirus infecting cucurbits in Sudan. *Arch. Virol.* 161, 2913-2915.
- Malandraki, I., Vassilakos, N., Xanthis, C., Kontosfiris, G., Katis, N.I. and Varveri, C., 2014. First Report of Moroccan watermelon mosaic virus in Zucchini Crops in Greece. *Plant Dis.* 98, 702.
- Martin, D. and Rybicki, E., (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16, 562-563.
- Martin, D.P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* 1(1), 1-5.
- Maynard Smith, J., 1992. Analyzing the mosaic structure of genes. *J. Mol. Evol.* 34, 126-129.
- Mohammed, H., Mangli, A., Zicca, S., El Hussein, A., Mohamed, M. and Tomassoli, L., 2012. First report of Papaya ringspot virus in pumpkin in Sudan. *New Disease Reports* 26, 26.
- Moury, B., Desbiez, C., Jacquemond, M. and Lecoq, H., 2006. Genetic diversity of plant virus populations: towards hypothesis testing in molecular epidemiology. *Adv. Virus Res.* 67, 49-87.
- Olarte-Castillo, X.A., Fermin, G., Tabima, J., Rojas, Y., Tennant, P.F., Fuchs, M., Sierra, R., Bernala, A.J. and Restrepo, S., 2011. Phylogeography and molecular epidemiology of *Papaya ringspot virus*. *Virus Res.* 159, 132-140.

- Owolabi, A.T., Rabenstein, F., Taiwo, M.A., Ehrigh, F. and Maiss, E., 2011. Characterization and identification of a potyvirus causing mosaic disease of *Cucurbita moschata* Duch ex. Poir in Calabar, South East Nigeria. *Int. J. Plant Pathol.* 2, 165-176.
- Padidam, M., Sawyer, S. and Fauquet, C., 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* 265, 218-225.
- Pirone, T.P. and Blanc, S., 1996. Helper-dependent vector transmission of plant viruses. *Ann. Rev. Phytopathol.* 34, 227-247.
- Posada, D. and Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc. Natl Acad. Sci. USA* 98, 13757-13762.
- Quiot-Douine, L., Lecoq, H., Quiot, J.-B., Pitrat, M. and Labonne, G., 1990. Serological and biological variability of virus isolates related to strains of papaya ringspot virus. *Phytopathology* 80, 256-263.
- Ray, S.C., 1998. SimPlot for Windows (version 1.6) Baltimore, Md. Distributed by author. <http://www.welch.jhu.edu/sray/download>.
- Revers, F. and Garcia, J.A., 2015. Molecular biology of potyviruses. *Adv. Virus Res.* 92, 101-199.
- Romay, G., Lecoq, H. and Desbiez, C., 2014. Zucchini tigre mosaic virus is a distinct species into the papaya ringspot virus cluster: Molecular and biological insights. *Arch. Virol.* 159, 277-289.
- Salminen, M.O., Carr, J.K., Burke, D.S. and McCutchan, F.E., 1995. Identification of Breakpoints in Intergenotypic Recombinants of HIV Type 1 by Bootscanning. *AIDS Res. and Human Retroviruses* 11, 1423-1425.
- Spetz, C., Taboada, A.M., Darwich, S., Ramsell, J., Salazar, L.F. and Valkonen, J.P.T., 2003. Molecular resolution of a complex of potyviruses infecting solanaceous crops at the centre of origin in Peru. *J. Gen. Virol.* 84, 2565-2578.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. and Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30, 2725-2729.
- Tomassoli, L., Tiberini, A. and Meneghini, M., 2010. Zucchini Yellow Fleck Virus is an Emergent Virus on Melon in Sicily (Italy). *J. Phytopathol.* 158, 314-316.

- Valli, A., Lopez-Moya, J.-J. and Garcia, J.A., 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. *J. Gen. Virol.* 88, 1016-1028.
- Yakoubi, S., Desbiez, C., Fakhfakh, H., Wipf-Scheibel, C., Marrakchi, M. and Lecoq, H., 2008a. Biological characterization and complete nucleotide sequence of a Tunisian isolate of Moroccan watermelon mosaic virus. *Arch. Virol.* 153, 775-781.
- Yakoubi, S., Lecoq, H. and Desbiez, C., 2008b. Algerian watermelon mosaic virus (AWMV): a new potyvirus species in the PRSV cluster. *Virus Genes* 37, 103-109.

## Figure legends

Fig. 1 (a) phylogenetic relationships along the whole nucleotide sequence of Su94-54 and reference isolates of six tentative or acknowledged species in the PRSV cluster, as determined with SplitsTree, and showing a network-like relationship between Su94-54, Su03-07 and MWMV suggestive of recombination., b) Nucleotide sequence identity along the genome of Su94-54 and six reference isolates in the PRSV cluster, as determined with SimPlot. Uncorrected sequence identities were calculated on sliding windows of 500 nt, with a step of 50 nt.

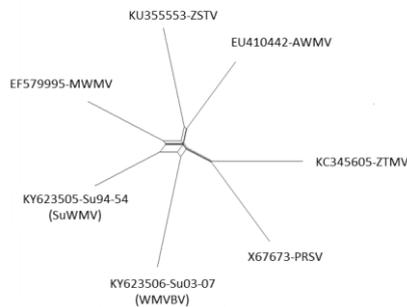


Fig. 1a

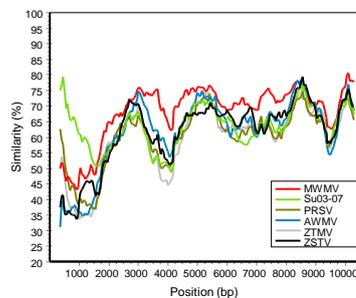


Fig. 1b

Fig. 2: Distance tree obtained for the nucleotide sequence of the core CP coding region of 45 members of the PRSV cluster, including the 2 new tentative species WMMVBV and SuWMMV, and three outgroup cucurbit-infecting potyviruses. Bootstrap values (500

bootstraps) above 70% are indicated for each node. Virus isolates originating from Africa are framed. The scale bar represents a genetic distance of 0.1.

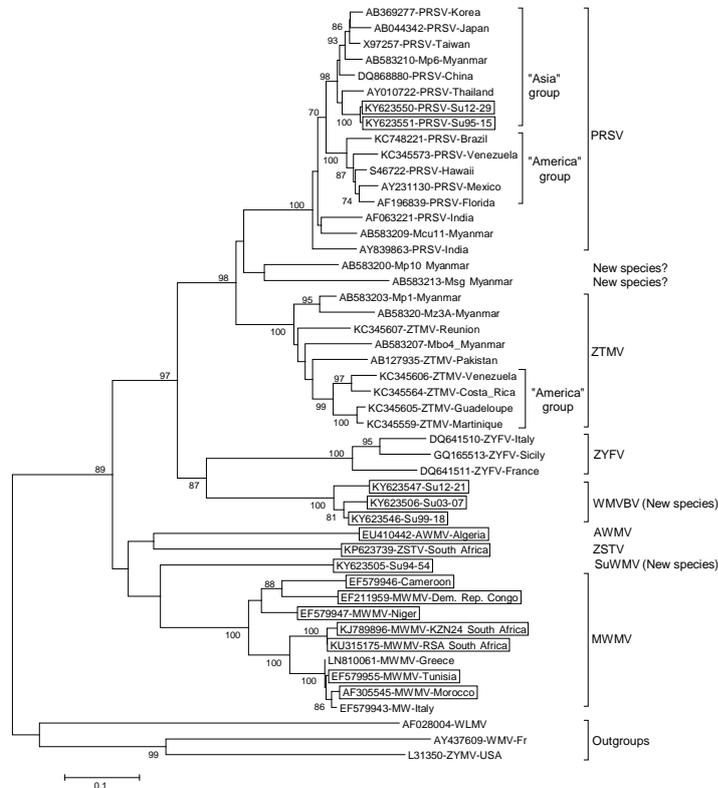


Fig. 3: Distance tree obtained for partial Nib-CP nucleotide sequence of SuWMV and MWMV isolates. All SuWMV isolates were from Sudan, whereas the geographic origin of the MWMV isolates is indicated on the tree. Bootstrap values (500 bootstraps) above 70% are indicated for each node. The scale bar represents a genetic distance of 0.05.



Table 1: Reaction of some differential hosts to viruses belonging to the 'PRSV cluster'

Differential hosts	Viruses				
	PRSV	MWMV	ZTMV	Su94-54 (SuWMV)	Su03-07 (WMVBV)
<i>Cucumis melo</i> 'Védrantais'	0/M	LL/TN	LL/SNS	0/M	LL/TN
<i>Cucumis melo</i> 'Ouzbèque'	0/M	0/M	0/M	0/M	0/M
<i>Cucumis melo</i> 'WMR29'	0/0(ns)	0/0(ns)	0/0	0/0(ns)	0/0
<i>Cucurbita pepo</i> 'Diamant F1)	M	M	MVB	M	SCS
<i>Chenopodium amaranticolor</i>	0/0	LL/0	0/0	0/0	LL/0

Left: inoculated leaf; right: systemic infection. 0: No symptom and no virus detected by DAS-ELISA (PRSV, MWMV, ZTMV and Su84-54) or by RT-PCR (Su03-07); M: mosaic; LL: local lesions; TN: lethal top necrosis; SNS: systemic necrotic spots; ns: occasionally a few systemic necrotic spots; VB: vein banding; SCS: systemic chlorotic spots.