Experimental evolution of cowpea mild mottle virus reveals recombination-driven reduction in virulence accompanied by increases in diversity and viral fitness

Zanardo, L.G.; Trindade, T.A.; Mar, T.B.; Barbosa, T.M.C.; Milanesi, D.F.; Alves, M.S.; Lima, R.R.P.N.; Zerbini, F.M.; Janssen, A.; Mizubuti, E.S.G.; Elliot, S.L.; Carvalho, C.M.

DOI
10.1016/j.virusres.2021.198389

Publication date
2021

Document Version
Final published version

Published in
Virus research

Virus Research 303 (2021) 198389

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Experimental evolution of cowpea mild mottle virus reveals recombination-driven reduction in virulence accompanied by increases in diversity and viral fitness

Larissa G. Zanardo [a], Tiago A. Trindade [b], Talita B. Mar [a], Tarsiane M.C. Barbosa [a], Diogo F. Milanesi [a], Murilo S. Alves [c], Roberta R.P.N. Lima [a], F. Murilo Zerbini [a], Arne Janssen [b,d], Eduardo S.G. Mizubuti [a], Simon L. Elliot [b], Claudine M. Carvalho [a,*]

[a] *Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, Brazil*
[b] *Departamento de Entomologia, Universidade Federal de Viçosa, Viçosa, MG, Brazil*
[c] *Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, CE, Brazil*
[d] *IBED, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands*

ARTICLE INFO

Keywords:
RNA virus
Virus evolution
Recombination
Evolution of virulence
Carlavirus
Bemisia tabaci

ABSTRACT

Major themes in pathogen evolution are emergence, evolution of virulence, host adaptation and the processes that underlie them. RNA viruses are of particular interest due to their rapid evolution. The *in vivo* molecular evolution of an RNA plant virus was demonstrated here using a necrotic isolate of cowpea mild mottle virus (CPMMV) and a susceptible soybean genotype submitted to serial inoculations. We show that the virus lost the capacity to cause necrosis after six passages through the host plant. When a severe bottleneck was imposed, virulence reduction occurred in the second passage. The change to milder symptoms had fitness benefits for the virus (higher RNA accumulation) and for its vector, the whitefly *Bemisia tabaci*. Genetic polymorphisms were highest in ORF1 (viral replicase) and were independent of the symptom pattern. Recombination was a major contributor to this diversity – even with the strong genetic bottleneck, recombination events and hot spots were detected within ORF1. Virulence reduction was associated with different sites in ORF1 associated to recombination events in both experiments. Overall, the results demonstrate that the reduction in virulence was a consequence of the emergence of new variants, driven by recombination. Besides providing details of the evolutionary mechanisms behind a reduction in virulence and its effect under viral and vector fitness, we propose that this recombination-driven switch in virulence allows the pathogen to rapidly adapt to a new host and, potentially, switch back.

1. Introduction

The evolution of virulence has become a major theme in evolutionary ecology and much theory has been developed, mainly based on a presumed trade-off between virulence and transmission (Alizon et al., 2009; Anderson, May, 1982). It has also been the subject of a number of studies in experimental evolution, which usually support the idea of the existence of an optimal level of virulence (Alizon et al., 2009; Froissart et al., 2010). Where virulence is initially high, it may be maladaptive, hence the parasite's evolution towards lower virulence; in fact, the pattern of selection on virulence is expected to change with time (Berngruber et al., 2013).

A number of experimental evolution studies have been conducted with RNA viruses (Bedhomme et al., 2012; Elena, 2016; Grubaugh, Andersen 2017; Koh *et al.*, 2017) because variability in such viral populations tends to be high. RNA viruses replicate rapidly, have short generation times and large effective population sizes that–associated with the lack of proofreading activity of RNA polymerases and their consequent high mutation and recombination rates–can generate highly polymorphic populations (Lalic et al., 2011). These factors can lead to rapid within-host evolution (Cuevas et al., 2016), adaptation to new niches and hosts (Koh et al., 2017), breakdown of host resistance (Hajimorad et al., 2005) and changes in virulence (Hasiów-Jaroszewska et al., 2011; Tribodet et al., 2005). High mutation rates, in association

* Corresponding author.
E-mail address: claudine.carvalho@ufv.br (C.M. Carvalho).

https://doi.org/10.1016/j.virusres.2021.198389
Received 27 November 2020; Received in revised form 3 March 2021; Accepted 9 March 2021
Available online 11 March 2021

with recombination events, offer a range of possibilities that allow viral populations to explore a diverse genotypic landscape and quickly end up at the fittest genotype (Burch, Chao, 2000). Recombination contributes by creating potentially advantageous genotypes and removing deleterious mutations, boosting the process of emergence. Rates of recombination depend on the prevalence of mixed infections with different species or strains, and also the ability of the viral replicases to undergo template switching. In summary, the maintenance of genetic variation is highly dependent on mutation and recombination rates, on the distribution of mutational effects on viral fitness and on the strength of selection and genetic drift within populations (Elena et al., 2011).

Cowpea mild mottle virus (CPMMV, family *Betaflexiviridae*, genus *Carlavirus*) is a single-stranded positive-sense RNA virus of ~ 8.2 Kb (Menzel et al., 2010; Zanardo et al., 2014a), which poses a major threat to soybean (*Glycine max*) production in Brazil as it causes the disease soybean stem necrosis. As the name implies, this disease can be highly virulent – the necrosis is systemic and can lead to plant death in 14 days (Zanardo et al., 2014a). However, there is considerable variation in disease symptoms, ranging from systemic necrosis to mild mosaic or even asymptomatic infection in the same soybean field (Almeida et al., 2003; Zanardo et al., 2014a, 2014b). While there is clearly a genetic basis for the variation in virulence observed in this system, there has been little indication in the studies conducted to date as to the underlying reasons for this variation (Zanardo et al., 2014b).

CPMMV is unusual among carlaviruses in that it is vectored (in a non-persistent manner) by the whitefly *Bemisia tabaci* MEAM 1 (Iwaki et al., 1982; Marubayashi et al., 2010; Naidu et al., 1998) rather than aphids. This whitefly is itself an important invasive pest (De Barro et al., 2011) and is notorious for its polyphagy (the ability to feeding on a wide variety of plant species). This polyphagy is understood to be an important factor contributing to the emergence and diversification of plant viruses (García-Arenal and Zerbini, 2019) and may well be important for CPMMV evolution. It seems likely that CPMMV has had to adapt to a variety of different plant hosts, which would be expected to contribute considerably to pathogen diversity. In this context, a suitable hypothesis is that the extremely high virulence of some CPMMV isolates (that are necrogenic) to some host genotypes is essentially maladaptive. A corollary to this hypothesis is that less aggressive (mild) isolates would be favored when subjected to selection in a given host.

Here, we used a highly virulent isolate of CPMMV and a very susceptible soybean cultivar to conduct an experimental evolution study. We sought to test if high virulence is a maladaptive trait that would be replaced over time by lower and more adaptive virulence in viral populations, and also which evolutionary mechanism(s) and viral genomic region(s) might be associated with this phenomenon. For this, we carried out successive inoculations of a necrogenic CPMMV isolate in susceptible soybean plants (cv. CD206). We observed that after six passages the highly virulent isolate that initially was causing infections with systemic necrosis symptoms switched to a less virulent isolate that caused infections with mild symptoms. The same was observed, although over fewer generations, from the infection of soybean plants with inoculum from a single local lesion caused by the same isolate in a *Nicotiana benthamiana* plant. Additionally, the reduction in virulence was accompanied by an increase in both viral fitness and fitness of its vector (the whitefly *Bemisia tabaci*), suggesting that necrogenesis is a maladaptive trait in CPMMV infections in this soybean cultivar. Genetic analyses showed that the reduction in virulence of the viral isolate was associated with mutations in different sites in ORF1 (viral replicase) located in recombinant blocks.

2. Material and methods

2.1. Experimental evolution– experimental design

Two experimental evolution experiments were conducted using a single isolate of CPMMV–CPMMV:BR:GO:10:5 (KC884249)–collected in January 2010 in a soybean field with plants exhibiting necrosis symptoms, in Cristalina, state of Goiás, Brazil (Zanardo et al., 2014a). The isolate was kept frozen at −80 °C. We were careful to use leaves from the original plant sample collected in the field for the first inoculation, carried out in both experiments, in order to ensure that all genetic changes in the viral isolate occurred during the performance of the experiments. Isolate CPMMV:BR:GO:10:5 is highly virulent and causes systemic necrosis; it is capable of causing death of the host plant in a period of 14 days post-inoculation (dpi). Necrotic symptoms are so intense that, regardless of the period in which the infection occurs, the plant host is unable even to produce new leaves, reinforcing how much necrogenesis could be a maladaptive trait in CPMMV infections. Here, an association between symptoms and virulence was used in which systemic necrosis was considered the highest level of virulence in CPMMV isolates. In the first experiment (single-host experiment), sampled leaves were used to mechanically inoculate three plants of soybean cv. CD206 at the VC stage. This soybean cultivar is very susceptible to CPMMV and inoculation always results in symptomatic infections. Inoculum was obtained by grinding the material in 0.1 M phosphate buffer, pH 7.2, with 0.1 % sodium sulfite. At 14 dpi, one of the plants (first passage–P1) was used to inoculate three more plants (second passage–P2) (Fig. 1a). Each of these three plants was then used to inoculate twelve more plants (third passage–P3) and each group of 12 plants was then considered a separate lineage (a, b and c; Fig. 1a). The plants were collected after 14 dpi, pooled and used for inoculation of 12 new plants (P4) for each lineage, and the procedure was repeated until P6 (Fig. 1a).

The inoculum concentration was not determined during the successive inoculations. Nevertheless, to ensure that the inoculum concentration was not interfering with the observed results, we used local lesions (which originate from 1 to 3 viral particles; Walker and Pirone (1972), as inoculum to conduct a second experiment (bottleneck experiment) (Fig. 1c). For this, *N. bethamiana* plants were mechanically inoculated as above using the original soybean samples of the isolate CPMMV:BR:GO:10:5 isolate, and local lesions were excised and used to inoculate three soybean plants (as above), one lesion per plant. One of these three plants (BP1 – first bottleneck passage) was used to inoculate three new soybean plants (BP2 – second bottleneck passage) (Fig. 1). All inoculated plants were maintained in gauze cages within a greenhouse (mean daily temperature 28 ± 2 °C) and were sprayed weekly with insecticides.

2.2. Real time PCR quantification of viral accumulation (single-host experiment)

RNA from healthy plants and from P1, P3 and P6 of the single-host experiment was extracted using the RNeasy Plant Mini Kit (Qiagen) at 14 dpi for all samples. The synthesis of cDNA was conducted with 500 ng of total RNA and the ORF6R primer (Zanardo et al., 2014a) using Superscript III reverse transcriptase (Invitrogen), according to the manufacturer's protocol. A standard curve was obtained using the CPMMV CP gene (867 nt), which was amplified from a CPMMV infectious clone (Carvalho et al., 2017). Serial dilutions of the CP gene PCR were performed in order to obtain concentrations corresponding to different numbers of copies (1,000–10 million copies). Sample quantification reactions were prepared with 50 ng of cDNA, specific primers for the CPMMV CP coding region (CPMMV-CPF, 5′ -ATA GTG AGA TGG CTG ATA AAC AAA AAC-3′, and CPMMV-CPR, 5' -TTC AGC ATC AAT GTC TGG AAG-3') which amplify a 79 bp fragment and SYBR Green PCR Master Mix (Applied Biosystems). Cycling conditions were 10 min at 94 °C, 40 cycles of 15 s at 94 °C and 1 min at 60 °C. A generalized linear model with a Gaussian error distribution (identity link) was used to compare average numbers of copies among treatments. Contrasts were determined with general linear hypothesis testing [function glht of the package lsmeans (Lenth, 2016) in R 2.15 (R Development Core Team, 2014)].

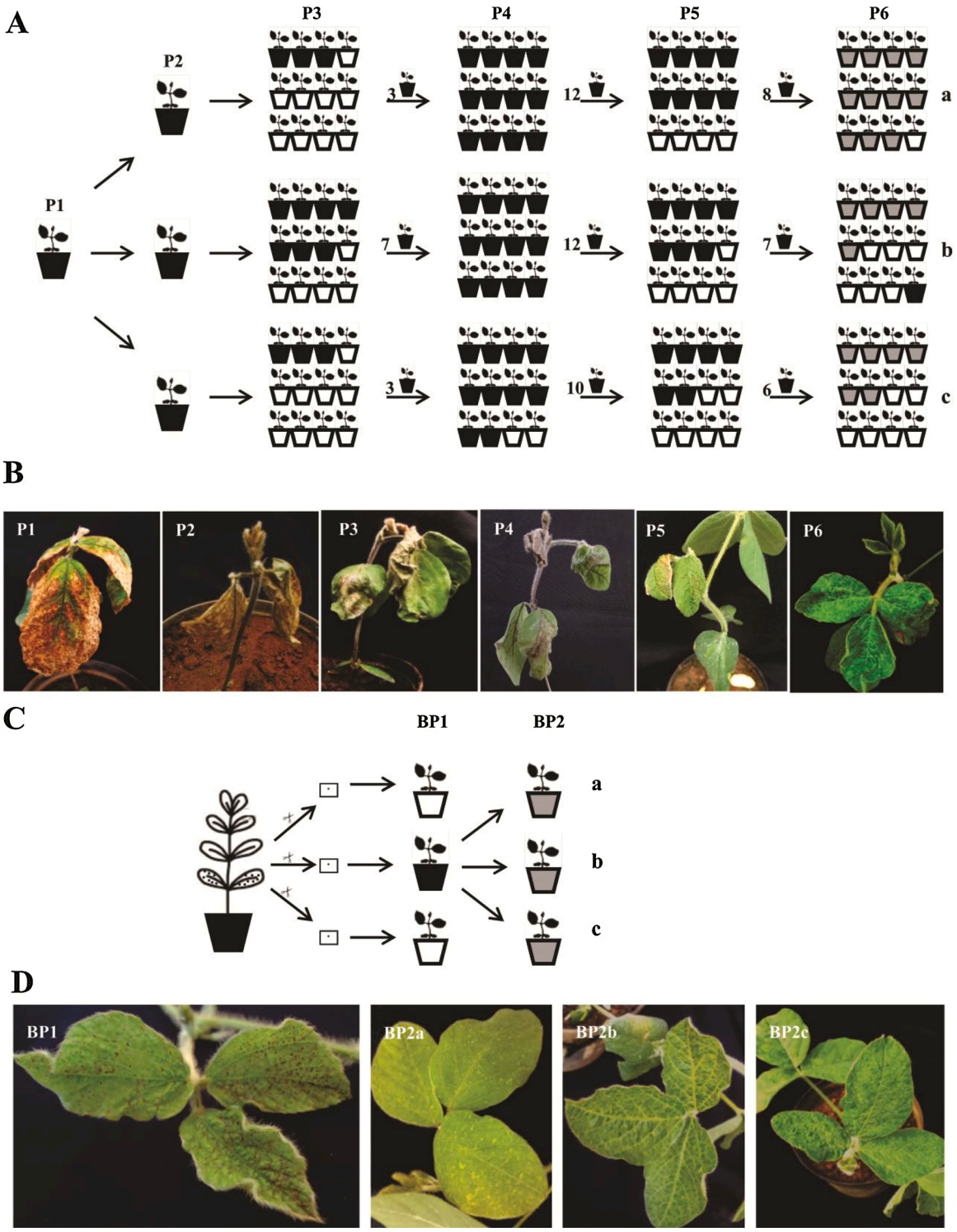


Fig. 1. Experimental design and phenotypic results of the two selection experiments (single-host and bottleneck) using isolate BR:GO:10:5 of *Cowpea mild mottle virus* (CPMMV) in soybean cv. CD206 at the VC stage (14 days of development) and the symptoms presented (14 days post-inoculation) in each passage. (A) Experimental design for the single-host experiment, showing the number of plants used at each step (passages P1-6) and the creation of three separate selection lineages (a, b and c). Numbers shown between the passages correspond to plant numbers used to inoculate the next plant generation. Black pots represent plants with symptoms of systemic necrosis, white pots are plants that showed no viral infection, and grey pots represent plants that showed mild symptoms. (B) Representative images of the symptoms observed. From P1 to P5, severe symptoms of systemic necrosis, bud burning and dwarfism, while in P6 symptoms were mosaics and vein clearing. (C) Schematic representation of serial inoculation of the bottleneck experiment with the same isolate excised from local lesions in *N. benthamiana* and re-inoculated in soybean (same cultivar and stage as the single-host experiment) in two passages (BP1-BP2). The letters a, b and c represent the replicates. (D) Symptoms observed over the serial inoculations on each passage after 14 days of inoculation. In BP1, necrotic symptoms were observed in the leaves, while in BP2 symptoms were very mild: mosaic, leaf deformation and vein clearing. The P1 image showed in B is the same from Zanardo et al. (2014a) (Fig. 1J).

2.3. Vector performance (single-host experiment)

Bemisia tabaci were obtained from a colony established in 2008 at the Federal University of Viçosa (UFV) and identified as *B. tabaci* Middle East-Asia Minor 1 (MEAM1) by sequencing the mitochondrial cytochrome oxidase I (mtCOI) gene (access numbers AF340215, AJ510071 and AJ510079). The insects were reared on cabbage plants (*Brassica oleracea* var. capitata) in gauze cages under greenhouse conditions. New soybean plants were inoculated with virus inoculum from P2 and P5 (same inoculum, but different plants from those used in the single host experiment), resulting in equivalent P3 and P6 generations. A separate isolate that causes mosaic symptoms in soybean cv. CD206 (CPMMV:BR:GO:01:1) (Zanardo et al., 2014a) and plants treated only with carborundum and buffer were used as positive and negative controls respectively.

Newly emerged whitefly nymphs (<24 h) were placed on P3 and P6 soybean plants (CPMMV-infected and uninfected controls) to evaluate survival and developmental times, at the V1 stage of soybean development. We used newly emerged nymphs, known as 'crawlers' as they are yet to settle and form scales, as they are easier to handle than subsequent instars. Adults were allowed to oviposit on healthy soybean plants for 24 h. After eggs hatched, ten nymphs were collected with a fine paint brush and placed on the abaxial surface of the first and second trifoliate leaves of 3–4 plants (replicates) resulting in the use of 30–40 nymphs. Nymphs that died after 24 h were replaced by others of the same age. The experiment was carried out in four randomized blocks in a room at $26 \pm 2\,^{\circ}C$ and a 12:12 photoperiod. Survival and development were recorded daily under a stereomicroscope until the emergence of adults. Dead nymphs were identified through their dark coloration and dried appearance. To eliminate handling effects on mortality, only nymphs that survived more than three days after placing them on the plants were used for survival analyses.

Survival and development data were first averaged per plant (replicate), and these averages were compared among treatments [P3, P6, C+ (soybean plants inoculated with CPMMV mosaic-inducing isolate); and healthy plants] using a linear mixed effects model (nlme package of R; Pinheiro et al., 2016). Blocks (A, B, C and D) were used as the random factor and treatments as the fixed factor. We started the analyses by including all factors and their interactions and by checking residual distributions. Non-significant interactions and factors were removed from the full model using the ANOVA function. Contrasts between treatments were assessed with the glht function (lsmeans package of R; Lenth, 2016).

2.4. RT-PCR and CPMMV cloning

Amplification and cloning of full viral genomes were performed from plants of all lineages from P1, P3 and P6 in the single-host experiment and from BP1 and BP2 in the bottleneck experiment (Fig. 1). The inoculum obtained from the original field sample was not amplified and sequenced, we wanted to standardize the host and the environmental conditions to which the virus would be submitted. Obtaining the viral genome sequences allowed us to investigate the genetic basis of symptom-induction in soybean plants. Total RNA was extracted from 100 mg of leaf tissue of individual or pooled soybean plants cv. CD206 (14 dpi) using the RNeasy Plant Mini Kit (QIAGEN). Reverse transcription (RT) was performed using Superscript III reverse transcriptase (Invitrogen), according to the manufacturer's protocol, using 500 ng of total RNA. PCR amplifications were carried out using Platinum High Fidelity DNA polymerase (Invitrogen). The primers and RT-PCR amplifications conditions were as described by Zanardo et al. (2014a), except for the 5′-end, for which the primer ORF1F (5′-GAA AAA CCC G-3′) and ORF1R (5′-TCT CGT TAG CTG AGG GTT-3′) were used. Amplified PCR products were gel-purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare), ligated into pGEM-T Easy Vector (Promega) and transformed into *Escherichia coli* DH5α cells. Ten different clones of each amplified region were selected. Plasmid DNA was purified using the Illustra plasmid Prep Mini Spin Kit (GE healthcare) and sequenced in both directions with universal primers (M13 F/M13R) by Macrogen (Seoul, South Korea).

2.5. Sequence assembly, identification and alignment

Sequences were assembled using DNA Baser Sequence Assembler v. 3.5 (Heracle Biosoft) and manually checked. Only those mutations found in more than one viral sequence were considered for subsequent analyses. The ORFs were located using ORF FINDER (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The nucleotide sequences were initially submitted to a BLAST search to check the sequence identification. Nucleotide sequences of each experiment were analyzed independently. Each ORF was aligned using the Muscle module (Edgar, 2004) in MEGA v. 5, based on amino acid sequences.

2.6. Description of molecular variability and divergence among sequences

Descriptors of molecular variability were estimated using the DnaSP software v. 5.10 (Rozas et al., 2003): (i) total number of segregating sites (s); (ii) G + C content; (iii) mean nucleotide differences between sequences (k); (iv) nucleotide diversity (π); (v) total number of mutations; (vi) number of haplotypes (h); (vii) haplotype diversity (Hd) and (viii) Watterson's estimator for the population-scaled mutation rate based on the total number of segregating sites (θ-W).

The π statistic was also calculated using a 100 nt sliding window, with a step size of 50 bases to estimate the nucleotide diversity throughout the section of the genome running from ORFs 1 to 6. Confidence intervals (CI) of 95 % were estimated for π values based on coalescent simulations (10,000 replicates) using DnaSP v. 5.10.

2.7. Bayesian phylogenetic analyses

Phylogenetic trees were constructed for each of the six ORFs individually (ORF1-6). Phylogenetic relationships were inferred using Bayesian inference (BI) in MrBayes v.3.0 (Ronquist, Huelsenbeck, 2003), with the evolution models selected by MrModeltest v.2.2 (Nylander, 2004) using the Akaike information criterion (AIC). Two runs with four Markov Chain Monte Carlo (MCMC) simulations were conducted simultaneously using 20 million generations, except for ORF1 of the first experiment (50 million generations) necessary for the MCMC chains to converge to a value less than 0.02, starting from random initial trees. Burn-in was set at 25 % from the resulting trees. Trees were visualized with Fig.Tree version 1.3.1 (http://tree.bio.ed.ac.uk/software/Fig.tree/) and midpoint rooted.

2.8. Discriminant analysis of principal components (DAPC)

Multivariate statistical analysis using Discriminant Analysis of Principal Components (DAPC) was performed in the package adegenet implemented in R software (Jombart, 2008) using nucleotide sequences from ORF1 of both experiments. The DAPC method is intended to describe the diversity between pre-defined groups of observations. We assumed that we had three different populations: population 1 (P1), population 2 (P3) and population 3 (P6) for the first experiment, and two different populations: population 1 (BP1) and population 2 (BP2) for the second experiment. The algorithm runs on a transformation of the raw data using Principal Component Analysis (PCA). All of the principal components were retained in order to conserve all of the variation in the original data. The key nucleotides involved in the cluster differentiation were determined and plotted using a threshold of 0.02 in single-host experiment and 0.002 in the bottleneck experiment in each discriminant function of DAPC.

2.9. Recombination analysis

Detection of potential recombinant sequences was performed using the Recombination Detection Program (RDP) v. 4.7 (Martin et al., 2010) incorporating the recombination detecting methods RDP (R), Geneconv (G), Maximum Chi Square (M), Bootscan (B), Sister Scan (S), Chimaera (C) and 3SEQ (3S). Default settings and a multiple comparison-corrected *P*-value cutoff of 0.05 were used throughout. Only those recombination events detected by more than four methods were considered reliable (Martin, 2009). The LDhat package in the RDP program was also used to estimate the recombination rate (ρ), to determine the recombination rates by site (ρ/site) and the relation of recombination rate with mutation rate (ρ/θ). A breakpoint density plot was constructed and the statistical significance of potential breakpoint hot- and cold-spots was tested using a permutation test with 10,000 permutations, a windows size of 200 and a step size of 100, with Confidence Intervals (CI) of 95 and 99 % in DnaSP v. 5.10. This test allows determination of whether the breakpoint distribution was significantly non-random.

2.10. Selection analysis

To test the occurrence of selection in populations, three types of neutrality tests were used: Tajima's D, Fu and Li's D* and F* using DnaSP v. 5.10. Gene and site-specific selection were also measured for each ORF and population in both experiments. Sites under negative and positive selection in each gene were determined using five different maximum likelihood-based algorithms: Single Likelihood Ancestor Counting (SLAC), Fixed Effects Likelihood (FEL), Internal Fixed Effects Likelihoods (IFEL), Random Effects Likelihood (REL) and Partitioning for Robust Inference of Selection (PARRIS) within the HyPhy software package (www.hyphy.org) with default conditions. The nucleotide substitution model applied was specific for each ORF and population. Phylogenetic trees corrected for recombination were inferred by GARD (available at the Datamonkey server) and used as input for the selection analysis.

2.11. Association test by phylogenetic approach and localization analysis

To determine the association between phenotypes (necrotic and mosaic) and the genotype changes observed in ORF1 after the passages, we conducted an association test and a localization analysis using the ALTree program (Bardel et al., 2006). For the association test, we determined the haplotypes in DnaSP v. 5.10, chosing the necrotic phenotype as control and the mosaic phenotype as case. Parsimony phylogenetic trees were inferred using PAUP v. 4.0 and rooted using the isolates CPMMV:P1:1 and CPMMV:BP1:1 from the first experiment and the bottleneck experiment, respectively. In the ALTree program all tests were done at phylogenetic tree level, and the *P*-value at each tree level was obtained by 10,000 permutation procedures.

To localize the sites associated with both phenotypes in ALTree, an *S* (supplementary) character was added for each affected (case) and unaffected (control) haplotype. After that, haplotypes including character *S* were reconstructed at ancestral nodes using PAUP v. 4.0. Susceptibility sites were localized through computing a correlated evolution index (Vi) calculated between each change of site and the changes of the character *S* in two possible directions of change. Sites whose evolution is most correlated with the *S* character are those most likely to be involved in the phenotype change.

3. Results

3.1. Serial passages lead to reduced virulence

We conducted serial inoculations (passages) of an isolate of CPMMV in soybean in three lineages in the single-host experiment (Fig. 1a). The numbers of plants that were successfully infected varied during the initial passages (P3-P5, Fig. 1a), and all infected plants suffered from necrosis (P1-P5), being much less extensive in P5 than in the previous passages (Fig. 1b). In the subsequent passage, P6, the phenotype changed simultaneously in all three lineages (a, b and c), with infected plants now showing mild symptoms characterized by mosaic and vein clearing, and only one plant displaying necrotic symptoms (Fig. 1a, b).

We chose to repeat the above experiment for the sake of validation, but also submitting the viral population to a genetic bottleneck to determine if inoculum concentration was the cause of phenotype variation. For this, we first inoculated plants of *N. benthamiana* to obtain local lesions, which served as inocula to infect soybean plants (a single lesion was used to provide inoculum to infect each soybean plant). At 40 dpi, one of these showed necrotic symptoms (Fig. 1c, d), although not as extensive as in most of the passages in the previous experiment. This single plant was used to infect three new soybean plants, as in the first experiment (Fig. 1c). These three soybean plants showed mosaic symptoms after 14 dpi – one generation (Fig. 1c, d), showing that composition of the inoculum could be associated with symptom variation.

3.2. Viral accumulation and vector performance increase once CPMMV becomes less virulent

Real-time PCR was used to determine viral accumulation and thus estimate within-host viral fitness in P1, P3 and P6 from the single-host experiment. For this, we assumed that amplification of the CP gene only occurred from intact viral particles. Viral accumulation differed significantly among plants from each passage (GLM, $F_{3,8} = 1462.4$, $P < 0.0001$) and was higher in P6 plants (with mosaic) than in P1 and P3 plants (with necrosis) (Fig. 2a), whereas there was no difference in accumulation between P1 and P3 (Fig. 2a). This analysis was not done in the Bottleneck experiment as there was only one plant per passage.

Viral infection on a plant can influence important components of an insect vector's fitness, such as survival, developmental time, fecundity or oviposition (Belliure et al., 2005; McKenzie, 2002; Jiu et al., 2007; Sidhu et al., 2009). We analyzed whitefly performance in P3 (necrotic), in P6 (mosaic), in a third treatment consisting of an isolate that causes mosaic symptoms naturally, and in a mock-inoculated control. There were significant differences in survival of whitefly nymphs to adulthood among treatments (LME, $\chi^2 = 24.2$, d.f. $= 3$, $P < 0.0001$; Fig. 2b). Survival on plants infected with the inoculum from P6 and with the mosaic strain did not differ significantly from one another or from the control, but survival on P3-infected plants was considerably lower: over 80 % of insects had died by day 5, *versus* less than 5% for the other treatments, and no whitefly reached adulthood on P3 plants. The time to develop from nymph to adult did not differ among other treatments (LME, $\chi^2 = 5.62$, d.f $= 3$, $P = 0.06$; Fig. 2c).

3.3. ORF 1 has high genetic polymorphism

Because the soybean genotype used for the serial passages was the same, it is reasonable to assume that genetic variations emerged in viral populations with higher fitness in both experiments. For the single-host experiment, we amplified 10 full genomes from each of three passages (P1, P3 and P6) and from the three distinct selection lineages in P3 and P6, resulting in the amplification of 70 full genomes. For the bottleneck experiment, 40 full genomes were amplified, 10 from each passage (BP1 and BP2 repetition a, b and c). The sequences of each lineage were pooled and analyzed together, resulting in three groups: P1, P3 and P6, with 10, 30 and 30 genomic sequences respectively, since the same symptom pattern was observed with only variations in the numbers of infected plants showing necrotic symptoms (Fig. 1a, c). The six typical carlavirus ORFs were detected in all genomes: ORF 1 encoding the putative viral replicase; ORFs 2, 3 and 4, encoding the triple gene block movement proteins (TGBp1, 2 and 3); ORF 5, encoding the coat protein (CP, 34 kDa); and ORF 6, encoding a cystein-rich, silencing suppressor

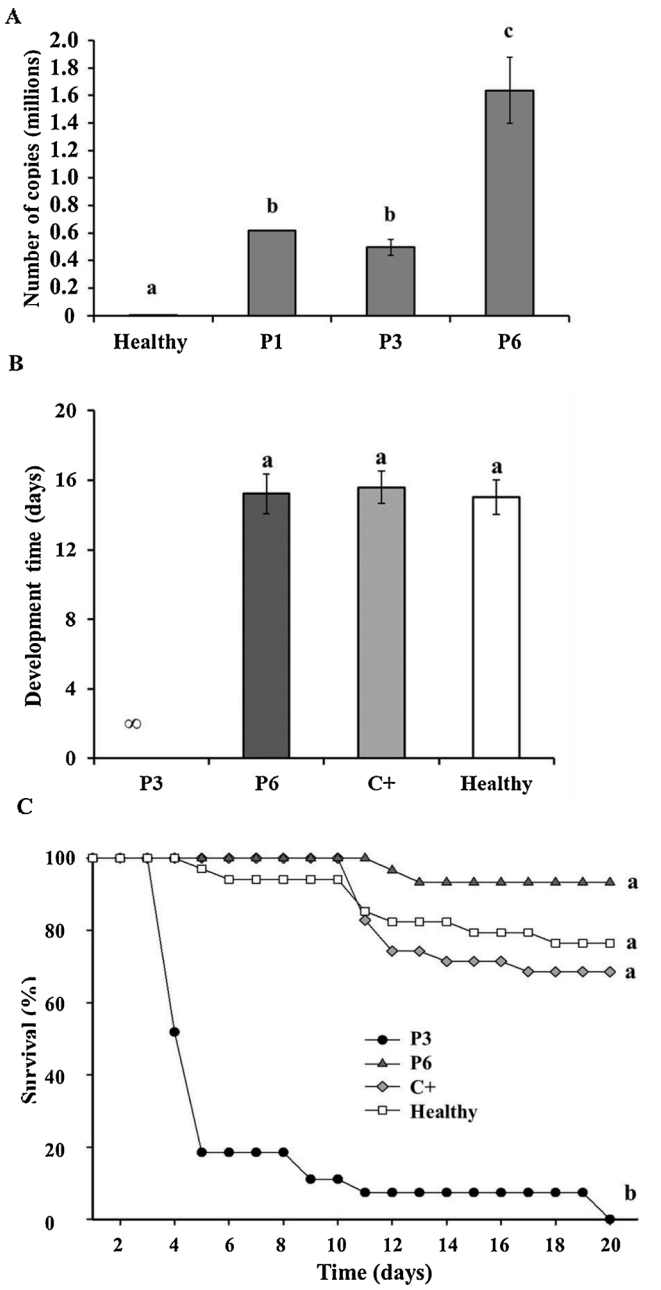


Fig. 2. Viral accumulation and vector fitness in plants following serial passages of CPMMV isolate BR:GO:10:5 in the single-host experiment. The isolate was passaged six times through soybean plants cv. CD206. P1 and P3 represent initial passages in which the virus caused systemic necrosis in the plants, while P6 represents a passage in which viral symptoms became milder (mosaic). (A) Viral accumulation in P1 (one plant), P3 (three plants per replication) and P6 passages (three plants per replication) after 14 dpi. Letters represent statistically significant difference ($P < 0.05$, contrasts after GLM with generalized linear hypothesis testing). Error bars are standard deviations. (B) Survival of whitefly (*Bemisia tabaci*) nymphs. P3, soybean plants infected with CPMMV BR:GO:10:5, necrotic variant; P6, soybean plants infected with CPMMV BR:GO:10:5, mild variant; C+, soybean plants infected with CPMMV BR:GO:01:1, a mosaic-inducing isolate; healthy, mock-inoculated plants. (C) Mean developmental times of newly emerged nymphs to adulthood on plants as under (B). ∞ means no data. The analyses in B and C were performed based on the average of three replications in which each treatment (C+, P3, P6 and healthy) consisted of 3-4 plants (replicates) with 30-40 nymphs.

protein (CRP) (Martelli et al., 2007; Zanardo et al., 2014a). The nt identity for all coding regions ranged from 93 to 100 % in the first experiment and from 98 to 100 % in the bottleneck experiment (Suppl. Figs. S1 and S2).

For both experiments, the genetic variability in all coding regions was evaluated (Table 1 and Suppl. Table S1). The sequences from each passage (P1, P3, P6, BP1 and BP2) were treated as populations and thus each experiment was analyzed separately. ORF1 was the most variable region for all populations in both experiments (Table 1 and Suppl. Table S1). The same G + C content was found for all regions of the genome, regardless of the generation evaluated. There was an increase in genetic polymorphism (nucleotide diversity - μ) with each passage in the single-host experiment (Table 1), but this was observed only for ORF1 in the bottleneck experiment (Suppl. Table S1).

In the first experiment, the mean pairwise number of nucleotide differences (π) for ORF1 increased threefold from P1 to P6 (Table 1). In the bottleneck experiment, ORF1 showed very small π values that were very similar for BP1 and BP2 (Suppl. Table S1). The population-scaled mutation rates (θ-w) ranged from 10^{-4} to 10^{-3} for ORFs 2–6, but for ORF1 this value was 10^{-2} for all passages in the single-host experiment (Table 1). For the bottleneck experiment, the values of θ-w were close to 10^{-3} for all ORFs, except for ORF3, for which it was 10^{-2} (Suppl. Table S1).

The π value for each ORF from each population was also determined in a 100 nt sliding windows (Fig. 3, Suppl. Fig. S3 and S4) and ORF1 was again the genomic region with the greatest variation, in both experiments (Fig. 3).

3.4. ORF1 is the region involved in population subdivision

We used Bayesian Inference (BI) in phylogenetic analyses to determine the genomic region which differentiated the populations. In the single-host experiment, the topology of the phylogenetic tree constructed based on ORF1 showed two clusters (Fig. 3c): the first cluster included all necrogenic isolates from P1 and P3 and 11 isolates from the P6 population. The second cluster only included isolates from P6 (Fig. 3c). The phylogenetic trees based on the other coding regions showed no such separation according to passage (Suppl. Fig. S5). Similar results were obtained in the bottleneck experiment (Fig. 3d and Suppl. Fig. S6). Thus, ORF1 is the genomic region that phylogenetically differentiated the P1, P3 and P6 populations.

Next, we used discriminant analysis of principal components (DAPC) to identify the sites in ORF1 that contributed most to differentiation among the populations in both experiments (Fig. 3e, f and Suppl. Fig. S7). A total of 73 sites were identified and described in the four regions (A–D) of the viral replicase (Fig. 3e and Suppl. Table S2) in the single-host experiment. In the bottleneck experiment, a total of 34 sites contributed most to differentiation between the two populations (Fig. 3f and Suppl. Table S2). No common site was found between experiments.

3.5. ORF1 is a recombination hotspot

Recombination events were only detected in ORF1: in the single-host experiment, 38 recombination events were detected (Suppl. Table S3); these occurred in all four variable regions (A, B, C and D) (Fig. 3e). In the bottleneck experiment only four recombination events were detected (Suppl. Table S4).

In the single-host experiment, eight of the 38 recombination events (events 13, 15, 16, 18, 25, 27, 29 and 37) were shared by more than ten isolates (Suppl. Table S3). Events 13, 15, 18, 25 and 37 were shared by isolates from the three passages (P1, P3 and P6). Of these eight events, three were of most interest: (a) event 15 was detected among isolates from P1, P3 and the 11 isolates from P6, (b) event 16 occurred in all isolates from P1 and P3 (necrotic isolates), involving the peptidase and helicase domains, and (c) event 29 was detected only in some isolates of P6 (Suppl. Table S3). Together, events 15 and 29 can explain the

Table 1
Descriptors of variability for *Cowpea mild mottle virus* (CPMMV) populations from the single-host experiment.

Genome region	Inoculation	Number of sequences	Region length	S^a	$G+C$	K^b	π^c	# of mutations[d]	H[e]	Hd[f]	θ-W[g]
ORF1 (REPLICASE)	P1	10	5577	413	0.401	96.956	0.01738 (± 0.00370)	413	10	1	0.02616
	P3	30		662	0.400	76.995	0.01380 (± 0.00172)	760	29	0.998	0.02995
	P6	30		788	0.411	285.391	0.05115 (± 0.00279)	1071	28	0.995	0.03565
ORF2 (TGB1)	P1	10	697	27	0.412	5.400	0.00776 (± 0.00299)	27	7	0.867	0.01371
	P3	30		5	0.411	0.395	0.00057 (± 0.00023)	5	6	0.310	0.00181
	P6	30		7	0.411	0.715	0.00103 (± 0.00028)	7	7	0.510	0.00254
ORF3 (TGB2)	P1	10	321	1	0.420	0.200	0.00062 (± 0.00048)	1	2	0.200	0.00110
	P3	30		6	0.421	0.400	0.00125 (± 0.00043)	6	7	0.366	0.00472
	P6	30		4	0.421	0.267	0.00083 (± 0.00036)	4	5	0.253	0.00315
ORF4 (TGB3)	P1	10	207	1	0.310	0.200	0.00097 (± 0.00074)	1	2	0.200	0.00171
	P3	30		3	0.310	0.262	0.00127 (± 0.00054)	3	4	0.251	0.00366
	P6	30		6	0.310	0.524	0.00253 (± 0.00090)	6	6	0.363	0.00732
ORF5 (CP)	P1	10	867	9	0.426	1.800	0.00208 (± 0.00042)	9	8	0.933	0.00367
	P3	30		14	0.426	1.239	0.00143 (± 0.00027)	14	12	0.777	0.00408
	P6	30		55	0.426	4.453	0.00514 (± 0.00299)	57	13	0.821	0.01601
ORF6 (NABP)	P1	10	342	5	0.418	1.000	0.00292 (± 0.00101)	5	5	0.667	0.00517
	P3	30		7	0.418	0.653	0.00191 (± 0.00056)	7	7	0.464	0.00517
	P6	30		8	0.419	0.772	0.00226 (± 0.00096)	8	5	0.409	0.00590

[a] Total number of segregating sites.
[b] Average number of nucleotide differences between sequences.
[c] Nucleotide diversity (with 95 % confidence intervals, using coalescent simulations).
[d] Total number of mutation.
[e] Haplotype number.
[f] Haplotype diversity.
[g] Watterson's estimate of the population-scaled mutation rate based on the total number of segregating sites.

clustering in the phylogenetic tree of 11 isolates from P6 (Fig. 3c, cluster 1). In the bottleneck experiment, event 1 only included isolates from BP1 (necrotic isolates) and event 2 only included isolates from BP2 (mild isolates) (Suppl. Table S4).

Recombination rates (ρ) were also determined for the passages (Table 2). The highest ρ values were in P6 and BP2, the two generations associated with changes in symptoms. Additionally, the recombination rate by site (ρ/site) was determined along the CPMMV nucleotide sequence (Table 2). The highest values of ρ/site were also in P6 and BP2 (Table 2).

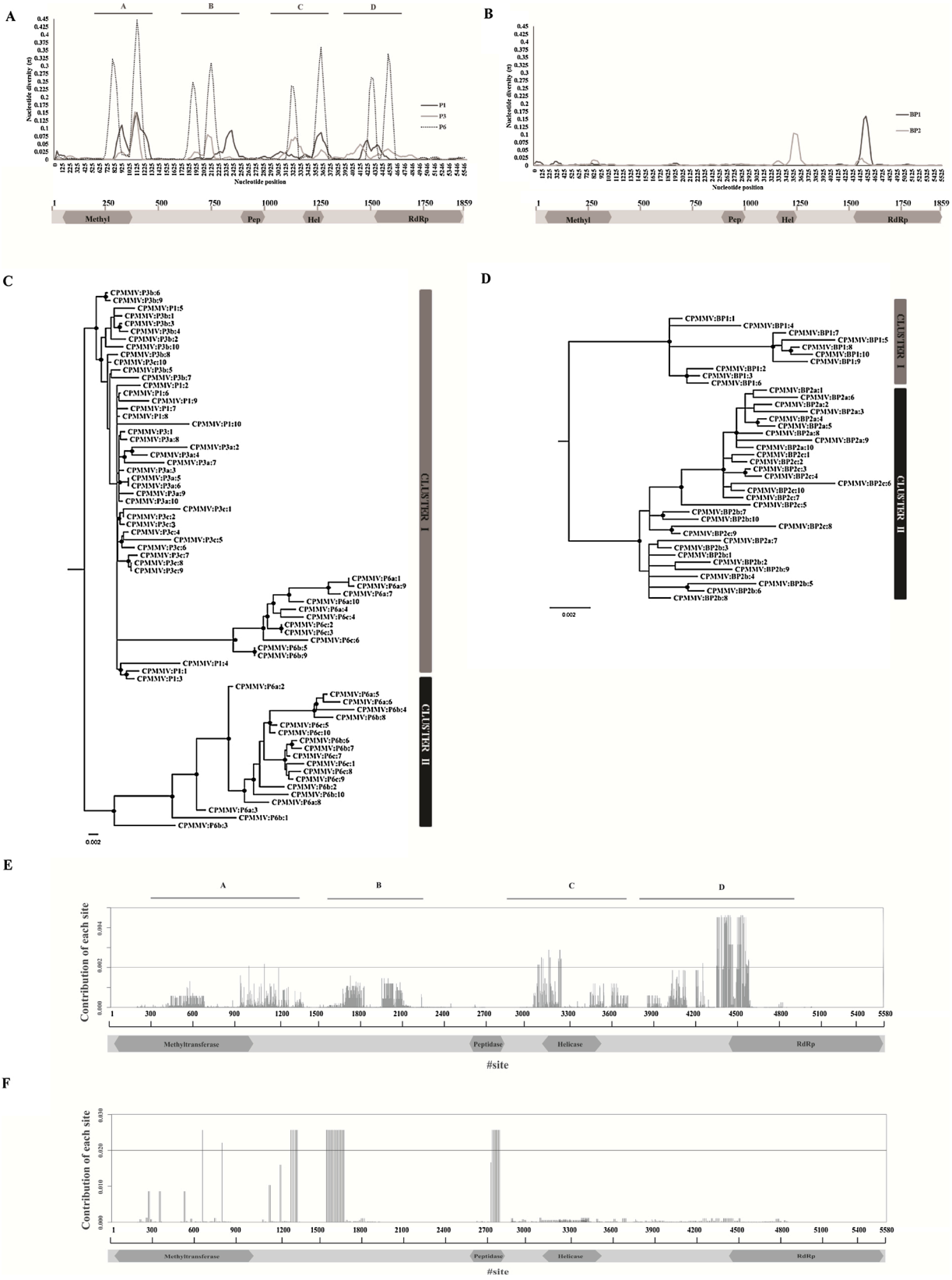
The ρ/θ ratio was determined to assess the relative contribution of recombination and mutation to the variability of CPMMV populations. For all populations, except P3, ρ/θ was higher than 1, suggesting that recombination contributed more to variability than mutations (Table 2). The importance of recombination was also confirmed by BI in phylogenetic analyses when recombinant blocks were excluded, the clustering previously observed in the ORF1 tree was disturbed (Suppl. Fig. S8).

A significantly higher density of breakpoints was observed in ORF1 when detectable breakpoint positions were plotted on a density map, suggesting that ORF1 is a recombination hotspot (Fig. 4). Together, these results suggest that recombination is responsible for most of the variation in ORF1.

3.6. Some codons are under positive selection in the replicase gene, but adaptive selection is not responsible for maintaining the diversity in ORF1

To evaluate how much of the genetic variation was being maintained by selection, analyses were performed on each CPMMV ORF using neutrality tests (Table 3 and S5). In the single-host experiment, all test values were negative, except for ORF1 in P6 (positive values). Statistically significant deviations from neutrality were observed for different populations and genomic regions (Table 3). Curiously, ORF5 was the only region where all values were negative and statistically significant (Table 3). In ORF1 of P6, all tests were non-significant, suggesting that the high polymorphism was not maintained by selection (Table 3). Similar results were observed in the bottleneck experiment for BP1 and BP2 (Table S5).

Site-by-site selection analyses were also performed for each genomic region using five algorithms: SLAC, FEL, IFEL, REL and PARRIS (Table 3 and S5). In the first experiment, evidence of negative (purifying) selection was observed for most codons under selection in ORF2−6 by all algorithms (Table 3 and S6). For ORF1 there were codons under positive selection, but many were also under negative selection, especially in the P6 population (Table 3 and S6). Strong evidence of positive selection in individual codons in ORF1 from P6 was detected for SLAC and PARRIS (Table 3 and S6). For the bottleneck experiment, the site-by-site selection analyses showed that the majority of the codons were subject to



(caption on next page)

Fig. 3. Sequences and phylogenetic analysis of ORF1 (viral replicase) of CPMMV isolate BR:GO:10:5 in two experimental evolution experiments (single-host and bottleneck) in soybean cv. CD206. (A) Average pairwise number of nucleotide differences per site (nucleotide diversity, π) calculated on a sliding window across ORF1. Sequences from a data set of 70 CPMMV sequences from the single-host experiment (dark grey line: ten sequences from P1; light gray line: 30 sequences from P3; dotted line: 30 sequences from P6. (B) Sequences from a data set of 40 CPMMV sequences from the bottleneck experiment (dark grey line: ten sequences from BP1; light gray line: 30 sequences from BP2. (C, D) Phylogenetic relationships of CPMMV isolates from the two experiments, using Bayesian inference (implemented in MrBayes 3.1). Support for the nodes is presented as filled circles (posterior probabilities from 0.95 to 1.0). Gray bars indicate the grouping of isolates causing severe (necrotic) symptoms, and black bars indicate isolates causing mild (mosaic) symptoms. (C) ORF1 tree using 70 sequences from the single-host experiment with the model GTR + I + G and 50 million generations; (D) ORF1 tree using 40 sequences from the bottleneck experiment with model GTR + I and 20 million generations; (C—D) The acronyms described in each branch of the tree represent each clone of each passage and lineage of both experiments. The number after the letter P corresponds to the passage and the letters a, b and c the lineages. As shown in Fig. 1. (E, F) Contribution of ORF1 sites to population subdivision in the single-host (E) and bottleneck (F) experiments, determined by discriminant analysis of principal components (DAPC). Heights of bars are proportional to the contribution of the corresponding site; bars above the threshold (0.02 or 0.002) correspond to sites that contributed most. The sites that were associated with the change of phenotype in association and localization tests can be checked in Suppl. Table S2. The letters A, B, C and D corresponded to the highly polymorphic regions shown in (A).

Table 2
Recombination rate (ρ) in populations of CPMMV from the single-host and bottleneck experiments.

Single-host experiment				
Pop.	ρ (CI)	ρ/site (CI)	θ/site	ρ/θ
P1	1166.679 (187.653–5048.163)	0.2125 (3.419×10^{-2}–0.919)	2.518×10^{-2}	8.441
P3	91.698 (62.221–131.29)	1.75×10^{-2} (1.187×10^{-2}–$2.505 \times 10^{-2)}$	2.774×10^{-2}	0.631
P6	3488.913 (282.606–8693.906)	0.6324 (5.122×10^{-2}–1.576)	2.525×10^{-2}	25.05
All	1237.834 (231.287–3265.553)	0.2244 (4.192×10^{-2}–0.5919)	4.165×10^{-2}	5.387
		Bottleneck experiment		
Pop.	ρ (CI)	ρ/site (CI)	θ/site	ρ/θ
BP1	957.927 (19.848–5935.066)	0.188 (3.899×10^{-3}–1.166)	4.791×10^{-3}	39.274
BP2	1635.841	0.318 (2.489×10^{-2}–0.879)	9.623×10^{-3}	33.065
All	9.375 (7.66–11.505)	2.112×10^{-3} (1.726×10^{-3}–2.592×10^{-3})	1.478×10^{-2}	0.143

*CI = Confidence intervals of 95 %; ρ= Recombination rate; θ= mutation rate.

negative selection in all ORFs from both populations (BP1 and BP2) (Table S7). Sites under positive selection were found in ORF1 (BP1 and 2) and in ORF6 in BP2 (Table S7). However, the evidence of positive selection was weak; the codons were detected only by IFEL and/or REL algorithms (Table S7).

The codons corresponding to the sites responsible for population subdivision were compared with the codons under selection and coincidences were only observed in the single-host experiment (Fig. 3E and F, Tables S2, S6 and S7).

Together, these results indicate that adaptive selection did not act to maintain the genetic variability of CPMMV populations.

3.7. Sites associated with phenotypic change are propagated across the extant population by recombination

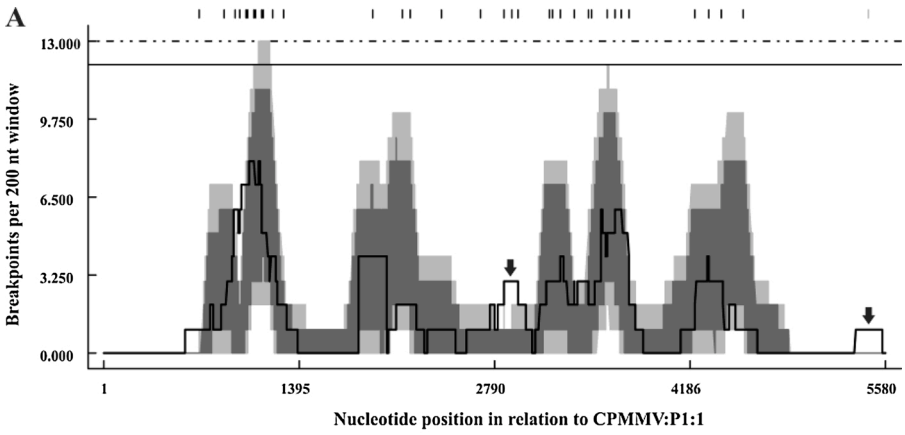
To try to associate the phenotypic change over successive inoculations with genotypic changes in ORF 1, association and localization tests were performed using the program ALTree. The 70 sequences generated in the first experiment were used in the analyses and 67 haplotypes were detected, 39 belonging to the control class (H01-H39, necrotic isolates) and 28 to the case class (H40-H67, mild isolates). In the association test the global *P*-value of the tree was significant ($P = 2 \times 10^{-4}$), indicating an association between phenotype and genotype changes (Suppl. Fig. S9).

A total of 27 sites were found involved in the phenotype change from necrosis to mosaic (Table 4) in the single-host experiment. These sites were compared with those sites responsible for population subdivision and selection: 15 and 17 coincidences were found, respectively (Table 4, Suppl. Tables S2 and S6). Additionally, most sites involved in phenotype change were located in recombinant blocks involving a number of isolates (events 13, 15, 16, 18, 20 and 25) and after recombination hotspot 1 (Figs. 4 and 5, Table 4 and Suppl. Table S3). This suggested that the recombination events are involved in the phenotype changes. The ALTree program uses a sequence of nucleotides to infer the association between sites and a given characteristic. For amino acid analysis, we carried out a manual and site-by-site analysis based on ALTree results. This analysis demonstrated that five mutations led to non-synonymous substitutions observed in all isolates from P6 (Table 4). Interestingly, another 11 mutations led to non-synonymous substitutions only in the isolates from P6 (Table 4) that formed a monophyletic branch (Cluster II) in the phylogenetic tree (Fig. 3C). All non-synonymous substitutions can be associated with recombination event 20.

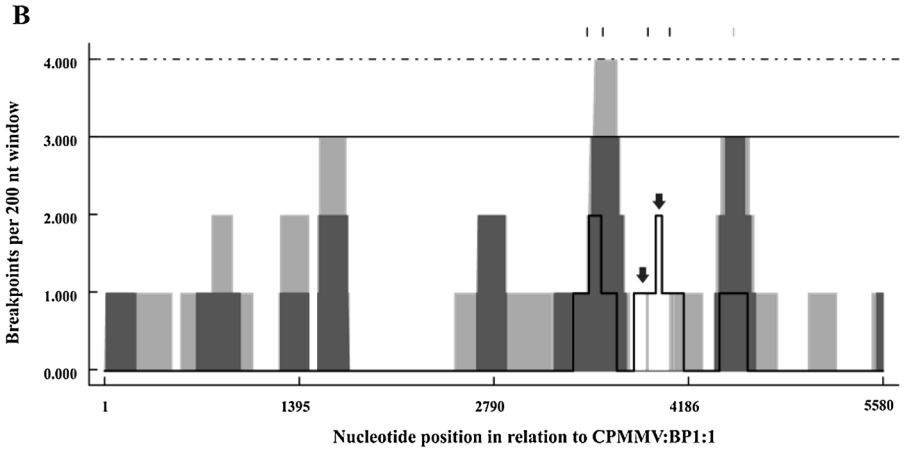
The association and localization tests were repeated with sequences from BP1 and BP2 in the bottleneck experiment. For the association test, 40 haplotypes were detected (from 40 sequences generated), 10 belonging to the control class (H01-H10, necrotic isolates) and 30 to the case class (H20-H40, mild isolates). The global *P*-value of the tree was significant ($P = 1 \times 10^{-4}$), indicating an association between phenotype and genotype changes also in the bottleneck experiment (Suppl. Fig. S10). In localization analyses, a total of 36 sites were associated with the phenotype change (Table 4). All the identified sites coincided with the sites responsible for population subdivision and none of them were under selection, but no coincidence was found with the single-host experiment (Table 4, Suppl. Tables S2 and S7). Many sites identified by the ALTree program were associated with the same codons and therefore the same amino acid substitution. The manual analysis of the mutations showed that there were 13 non-synonymous substitutions identified in the BP2 isolates in the genetic bottleneck experiment (Table 4). All the sites and non-synonymous substitutions were located the in recombinant block from event 1 in ORF1 (Fig. 5, Table 4 and Suppl. Table S4).

4. Discussion

We have shown that the high virulence of an isolate of CPMMV that causes systemic necrosis evolved to lower virulence, causing milder symptoms, in a consistent pattern across separate lineages and two independent experiments. Our study indicates that this could be due to a rapid evolutionary process driven primarily by recombination. Several sites in ORF1 located in recombinant blocks were associated with the change in symptoms. These alterations in the viral genome generated variants with higher fitness and also increased fitness of the insect vector.



#	Type	Position*	Region	Significance
1	Hot-spot	2811-3055	Pep-IG	Locally
2	Hot-spot	5366-5561	RdRp	Locally



#	Type	Position*	Region	Significance
1	Hot-spot	3797-3881	IG (Hel-RdRp)	Locally
2	Hot-spot	3899-4053	IG (Hel-RdRp)	Locally

Fig. 4. Distribution of recombination breakpoints within ORF1 of CPMMV isolate BR:GO:10:5 in two experimental evolution experiments (single-host and bottleneck) in soybean cv. CD206. (A) Recombination breakpoints in the single-host experiment. Sequences from a data set of 70 CPMMV sequences, including P1, P3 and P6 (first, third and sixth passages, respectively). (B) Recombination breakpoints in the bottleneck experiment. Sequences from a data set of 40 CPMMV sequences, including BP1 and BP2 (first and second passages, respectively). All detectable breakpoint positions are indicated by small vertical lines at the top of the graph. A 200 nucleotide window was moved along the alignment one nucleotide at a time and the number of breakpoints detected within the window region was counted and plotted (solid line). The upper and lower broken lines indicate 99 and 95 % confidence thresholds, respectively, for globally significant breakpoint clusters. Dark and light gray areas indicate local 95 and 99 % breakpoint clustering thresholds, respectively. Vertical black arrows indicate recombination hot spots.

Since isolates with both the severe and mild phenotypes can be found in soybean fields (Almeida et al., 2003; Zanardo et al., 2014a), there was an expectation that a change in virulence would occur over the course of the successive inoculations. Systemic necrosis is very damaging to the host plant, resulting in rapid death. As CPMMV and its whitefly vector depend upon their host for reproduction and transmission (or dispersal in the case of the insect vector) to new hosts, the high virulence associated with systemic necrosis may well be detrimental to the virus. Nevertheless, it was striking that the change in virulence was so abrupt.

The single-host experiment allowed us to examine both viral accumulation and vector performance and in particular allowed a comparison between populations of the same viral isolate before and after it lost its capacity to cause necrosis. The pattern of symptoms observed is directly related to viral accumulation: the more aggressive the symptoms, the lower the CPMMV accumulation. Real-time PCR has been used as a tool for determining viral fitness (Carrasco et al., 2007; Péréfarres et al., 2014), under the assumption that higher viral accumulation means higher fitness. This is because fitness can be defined as “the capacity of a virus to produce infectious progeny in a given environment” (replicative fitness) (Domingo and Holland, 1997). The greater replicative ability ensures the survival of the next generations and increases the chance of transmission by the whitefly. Thus, the lowest amount of virus particles after P1 and P3 was probably a direct consequence of a high fitness cost for the virus (Leach et al., 2001). This is comparable to what was observed for soybean mosaic virus in soybean, in which an increase in virulence resulted in reduced fitness (Khatabi et al., 2013). It is important to note that the change in symptoms caused by successive inoculations of the CPMMV necrogenic isolate was observed in the soybean cultivar CD206. This same isolate may be able to cause other symptoms in different soybean hosts and cultivars, which may or may not include necrosis. Thus, the behavior of the vector and the viral isolate may be different from that observed here during experiments.

It is possible that the maintenance of high virulence in nature would occur if the virus were transmitted during the early stages of the infection. Although nymphal stages of the whitefly are almost entirely sessile, adults are highly mobile and take flight at the slightest disturbance. As the virus is transmitted in a non-persistent fashion, this may facilitate transmission even if the virus is lethal to the plant, since the vector does not have to be in contact with the host for a long time. Meanwhile, less virulent isolates could be transmitted at early or late stages and these could even represent distinct strategies for transmission. It has been shown that a highly virulent isolate of cucumber mosaic virus (CMV) that caused necrosis in tomato required high transmission rates ensured by a high density of its aphid vector (Escriu et al., 2003). The same could be expected for CPMMV populations, with a high density of whiteflies being necessary to ensure that more virulent isolates are transmitted

Table 3
Neutrality tests and site-by-site selection analysis for each CPMMV in the single-host experiment, according to four algorithms implemented in the Datamonkey server.

Genome region	inoculation	Tajima'D	Fu & Li's D*	Fu & Li's F*	SLAC[a]		FEL[b]		IFEL[c]		REL[d]		PARRIS[e]
					N	P	N	P	N	P	N	P	P
	P1	−1.85041*	−2.04262*	−2.25531*	3	–	22	2	1	9	PS	–	–
ORF1	P3	2.33920***	−3.62413**	−3.77908**	4	–	55	7	9	34	–	69	–
	P6	0.17794ns	0.13339ns	0.17250ns	22	2	46	16	35	18	28	136	+
	P1	2.07497***	−2.45320**	−2.66135**	–	–	2	–	–	–	–	–	–
ORF2	P3	−1.87346*	−2.36062ns	−2.57755ns	–	–	1	–	–	1	–	–	–
	P6	−1.75911ns	−0.81656ns	−1.28327ns	–	–	2	–	–	–	PS	–	–
	P1	−1.11173ns	−1.24341ns	−1.34668ns	–	–	–	–	–	–	–	–	–
ORF3	P3	−2.09995*	−3.43598**	−3.53709**	–	–	–	–	–	–	PS	–	–
	P6	−1.88948*	−2.99378*	−3.10061*	–	–	–	–	–	–	PS	–	–
	P1	−1.11173ns	−1.24341ns	−1.34668ns	–	–	–	–	–	–	–	–	–
ORF4	P3	−1.53889ns	−1.47512ns	−1.72941ns	–	–	1	–	–	–	PS	–	–
	P6	−1.86605*	−1.88589ns	−2.18941ns	–	–	2	–	–	–	2	–	–
	P1	−1.88589*	−2.21874**	−2.40603**	–	–	1	–	–	–	–	1	–
ORF5	P3	−2.16704*	−2.63942*	−2.92188*	–	–	2	–	–	–	–	–	–
	P6	−2.58928***	−4.60532**	−4.65377**	1	–	7	–	2	–	1	–	–
	P1	−1.74110*	−2.01007ns	−2.17902ns	–	–	–	–	–	–	–	–	–
ORF6	P3	−1.86290*	−1.51321ns	−1.88745ns	–	–	1	–	–	–	PS	–	–
	P6	−1.87584*	−2.47845ns	−2.68001*	–	–	2	–	–	–	PS	–	–

(P) Sites under positive selection; (N) sites under negative selection; (PS) purifying selection; (-) no site under selection; (+) sites under positive selection.
* $p < 0.05$.
** $p < 0.02$(P).
*** $p < 0.001$.
ns non-sigficant.
[a] Single Likelihood Ancestor Counting (SLAC).
[b] Fixed Effects Likelihood; (FEL).
[c] Internal Fixed Effects Likelihood (IFEL).
[d] Random Effects Likelihood (REL).
[e] Partitioning for Robust Inference of Selection (PARRIS).

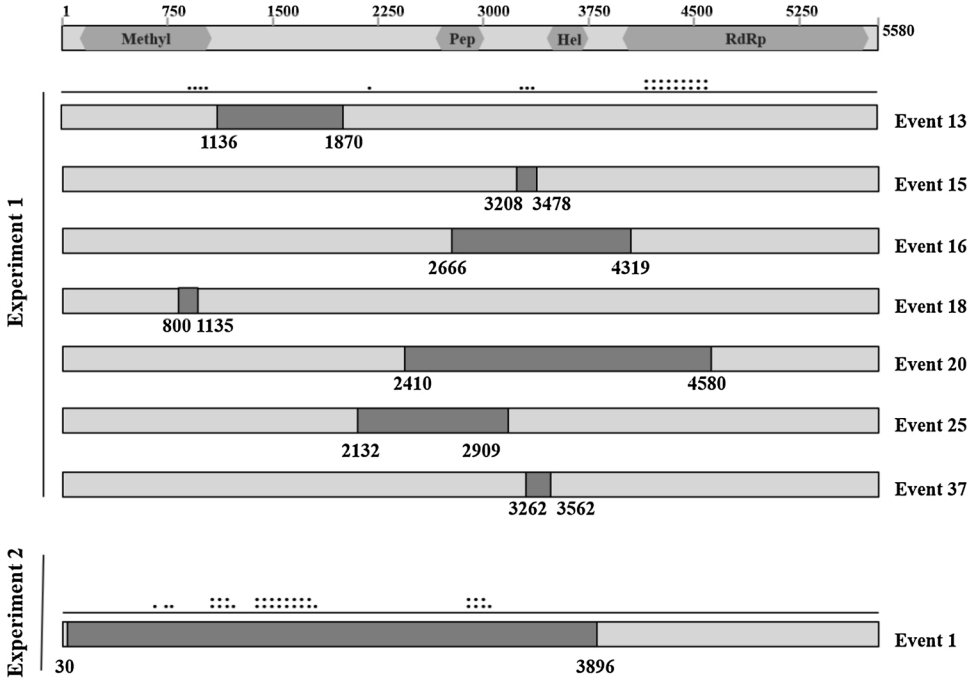


Fig. 5. Schematic representation of recombination blocks and sites in ORF1 associated with phenotypic changes in CPMMV isolate BR: GO:10:5 in two experimental evolution experiments (single-host and bottleneck) in soybean cv. CD206. (A) Recombination events detected in the single-host experiment in more than three sequences, and which coincide with the position of the sites involved in phenotype change. (B) Recombination event detected in the bottleneck experiment, coinciding with the position of the sites involved in phenotype change. Recombinant blocks are shown in dark gray, and were identified with the RDP program v. 4.7. The position of the sites associated with the phenotype changes are indicated by small dots in the upper portion of the figure.

under field conditions. In addition to the effects of infection on the plant host, viruses can also have indirect effects on the performance of vectors and ultimately this may affect transmission (Luan et al., 2013; Mauck et al., 2012; Mauck, 2016; Mauck et al., 2010; Wan et al., 2015; Zhang et al., 2012). The inability of the whitefly to survive and develop in a plant infected with the necrotic isolate is an obvious adaptive disadvantage compared to the non-necrotic symptoms induced by the mild variant. In addition, it suggests an ecological and evolutionary cost of inducing systemic necrosis for the virus. Even if adult vectors passing from plant to plant and probing may effect viral transmission, individuals that develop on an infected plant will also contribute considerably to transmission if they are able to develop to adulthood. Thus, it is advantageous for the maintenance of CPMMV in the field to induce mild symptoms. Another scenario that we should consider is that even though necrosis may be maladaptive to the virus, it is observed at around 10–12 dpi in soybean plants cv. CD206, and this period of development of the

Table 4
Description of sites in ORF1 of CPMMV isolates from both experiments as potential sites involved in phenotype changes.

Single-host Experiment				Bottleneck Experiment			
Site	Codon	nt substitution (aa)	Vi	Site	Codon	nt substitution (aa)	Vi
1082	344	A–>C (A/A)	6.24695	648	216	G–>A (K/K)	7.682213
1096	365*	T–>C (E/Q)	5.364934	819	273	C–>T (I/I)	7.682213
1119	373*	C–>A (C/R)	6.24695	957	319	C–>T (F/F)	4.287483
1122	374*	T–>G (S/R)	5.364934	1354	452	G–>A (E/K)	7.682213
1130	377*	G–>C (S/T)	5.364934	1357	453	A–>G (R/D)	7.682213
2345	782*	A–>G (G/D)	7.714678	1358	453	G–>A (R/D)	7.682213
3225	1075*	A–>C (G/G)	7.714678	1359	453	A–>T (R/D)	7.682213
3243	1081	G–>T (V/V)	6.24695	1360	454	T–>C (S/L)	7.682213
3246	1082*	A–>G (S/S)	7.714678	1361	454	C–>T (S/L)	7.682213
4463	1488*	G–>A (R/H)[a]	7.714678	1362	454	T–>C (S/L)	7.682213
4464	1488*	A–>C (R/H)[a]	11.00037	1627	543	G–>A (E/R)	7.682213
4465	1489*	C–>T (L/L)	11.00037	1628	543	A–>G (E/R)	7.682213
4468	1490	A–>G (R/G)[b]	11.00037	1629	543	G–>A (E/R)	7.682213
4470	1490	G–>A (R/G)[b]	11.00037	1630	544	A–>T (I/S)	7.682213
4482	1494*	A–>C (K/N)[b]	11.00037	1631	544	T–>C (I/S)	7.682213
4504	1502*	T–>C (L/L)[b]	7.714678	1632	544	C–>A (I/S)	7.682213
4514	1505	C–>A (A/Q)[b]	11.00037	1633	545	A–>G (S/V)	7.682213
4524	1508	T–>G (F/W)[a]	11.00037	1634	545	G–>T (S/V)	7.682213
4525	1509	G–>A (G/N)[b]	11.00037	1636	546	T–>C (S/L)	7.682213
4530	1510*	A–>T (R/H)[b]	11.00037	1637	546	C–>T (S/L)	7.682213
4546	1516	C–>A (L/I)[a]	7.714678	1638	546	T–>G (S/L)	7.682213
4548	1516	C–>T (L/I)[a]	11.00037	1639	547	G–>C (A/L)	7.682213
4565	1522*	T–>C (F/S)	11.00037	1640	547	C–>T (A/L)	7.682213
4567	1523*	C–>T (L/W)[b]	11.00037	1641	547	T–>A (A/L)	7.682213
4570	1524	G–>A (D/N)[b]	11.00037	1644	548	A–>T (K/N)	7.682213
4577	1526	T–>A (V/N)[b]	11.00037	1646	549	T–>C (F/S)	7.682213
4578	1526	G–>C (V/N)[b]	11.00037	1647	549	C–>A (F/S)	7.682213
				2770	924	C–>A (H/I)	4.287483
				2771	924	A–>T (H/I)	7.682213
				2772	924	T–>A (H/I)	7.682213
				2773	925	A–>C (T/H)	7.682213
				2774	925	C–>A (T/H)	7.682213
				2775	925	A–>T (T/H)	7.682213
				2776	926	T–>G (C/A)	7.682213
				2777	926	G–>C (C/A)	7.682213
				2778	926	C–>T (C/A)	7.682213

* Codons also submitted to positive selection. Vi = evolution index.
[a] Non-synonymous substitutions associated with all P6 isolates.
[b] Non-synonymous substitutions associated with P6 isolates from Cluster II in phylogenetic tree.

infection is sufficient for the virus to be acquired by the vector and transmitted to new hosts. In addition, this short period of time could be a way to ensure lower virus accumulation which in turn would increase the efficiency of transmission.

Determining the nature of the phenotypic change after successive passages was challenging. Changes in symptom patterns are usually associated with non-synonymous mutations in one or a few sites (Diveki et al., 2004; Hirata et al., 2003). However, in the case of CPMMV the phenotypic changes in the viral population were governed by a more complex interaction network involving a set of mutations in ORF1, which encodes the viral replicase. ORF1 was the most polymorphic region of the viral genome in both the single-host and the bottleneck experiments. The fact that ORF1 was the most variable region of the viral genome was somewhat surprising, considering that it encodes the only viral protein which is required for replication. One would assume that a protein involved in such a fundamental process would be under strong negative selection pressure, and thus would be less prone to variation. Obviously, in the context of our experiments, this assumption does not hold. In fact, our results indicate that interactions between the viral RdRp and host factors may be the genetic basis for the induction of necrotic *vs*. mild symptoms. If follows that non-synonymous substitutions in ORF1 are essential for phenotypic changes to occur: in the single-host experiment, 11 non-synonymous mutations in ORF1 could be associated to the phenotype change, while 13 different non-synonymous mutations were associated with the change in the bottleneck experiment. These substitutions were different between the two experiments, suggesting the occurrence of multiple determinism, which means that, although the phenotypic change occurred in both experiments, there are different mutations associated with the same event. This would occur due to the increase in genetic polymorphism in ORF1. High levels of variation are commonly found in the ORF1 of viruses from the *Betaflexiviridae* family in inter-host analyses (Alabi et al., 2010; Moles et al., 2007), but this is the first time that it has been observed in an intra-host analysis.

Genetic bottlenecks to which the initial inoculum was subjected in the bottleneck experiment clearly affected the genetic variability of the CPMMV population. This is probably because the drastic reduction in effective population size limited the possible number of viral genomes available to found the new viral population, thus accelerating the change in phenotype. This probably means that there is only one viral genotype that causes necrosis from the local lesion in *N. benthamiana*, so the change in symptom patterns is faster than in the single-host experiment, where there is a viral population with several genotypes that cause necrosis. We believe that there may already be viral genotypes that cause milder symptoms, but they are less prevalent in the viral population. The genetic bottleneck limits the fixation of beneficial mutations, but phenotypic changes were still observed. This suggests that even though the mutations are not fixed, the number of viral generations obtained within the host was sufficient to generate variation and maintain it in the viral population.

Although the changes in phenotype occurred earlier in the bottleneck experiment, these changes occurred in synchrony across lineages in each experiment. This suggests that the process was neither stochastic nor determined by inoculum concentration, but was determined by

evolutionary mechanisms that changed the inoculum composition (different genotypes were generated). Mutation rates could be very similar across passages (Table 1) and therefore phenotypic changes occur at the same time. In addition, data also suggests that since the mutation (or mutations) associated with changes has occurred, the variant genotype increases in the population and thus the phenotypic change is observed. This would also explain the synchronized change across lineages. In principle, the changes observed were not determined by one or a few mutations, but by a series of mutations (mechanism one). The easiest and fastest way to acquire these beneficial mutations is by recombination (Elena et al., 2011) (mechanism two). Strikingly, all mutations associated with phenotypic change were located in recombination hotspots in ORF1. Each mutation or set of mutations, resulting in non-synonymous substitutions, may be contributing to increased viral fitness by changing protein structure and thus affecting the interaction between the CPMMV replicase and host factors involved in triggering defense responses which culminate in systemic necrosis. As we found different mutations associated with more than one recombinant block, more than one genotype is able to cause the change in the phenotype. A single plant with symptoms of necrosis was found in the sixth passage, demonstrating that although the genotypes and consequently the phenotype has changed, there were still, albeit to a lesser extent, viral genotypes present that caused necrosis.

Further to the above, the ORF1 regions affected by mutation and recombination events were different in both experiments, suggesting that distinct changes in the primary structure of the protein will differentially affect its interaction with host factors, resulting in a different symptom pattern, since several mutations resulted in non-synonymous substitutions. Such epistatic effects can be positive or negative (Bernet and Elena, 2015; Lalic and Elena, 2015). The mutations detected in ORF1, associated with recombination events, led to higher viral accumulation and thus increased viral fitness. Thus, it is reasonable also to suggest that the effect on the phenotype could be a result of a positive epistatic effect. This hypothesis could be test in infectivity experiments in which different combinations of mutations are introduced in an infectious clone.

Although we focused our analyses on changes in the sequences that we were evaluating, we always sought to associate the variations in genotypes with the phenotype. The accumulation of defective interfering (DI) RNA viral particles can occur under certain conditions of temperature, inoculum concentration and host, and this may influence virulence and viral accumulation. This is because DIs are packaged as wild type RNAs and interfere with viral replication, replicating in a dependent and preferential way at the expense of the non-defective parental virus (Huang and Baltimore, 1970). No carlavirus described in the literature to date has been associated with DI RNAs, including studies specific for CPMMV in which metagenomics experiments were performed (Rosario et al., 2014). Furthermore, Inoue-Nagata et al. (1997) demonstrated that the formation of DIs upon serial inoculations of a tospovirus (up to 15 successive inoculations) is favored by high initial inoculum concentration and low temperature (16 °C; at 30 °C, no DIs were observed), which is in contrast with our experiments [high and low viral inoculum concentrations and high temperatures (28+/−2 °C)]. Thus, the appearance and accumulation of DIs is unlikely to be associated with the phenotype change observed in our experiments with CPMMV.

Our results indicate that the set of mutations were acquired by different variants of CPMMV through the infection and were propagated *via* recombination in ORF1 during serial plant inoculations. Having higher fitness, the frequency of these variants increased in the population. Extrapolating this situation to the field, new mild variants would easily arise and accumulate to higher levels within the plant. These variants would also be more beneficial for the insect vectors, resulting in their rapid selection since the vector would be able to survive longer and oviposit at a higher rate. It is possible, however, that this process might render the variant less adapted to other host genotypes. Thus, it would be interesting to determine if the new mild variant has higher fitness across a range of host genotypes. If fitness is decreased in certain genotypes, we may observe a reversal of the evolutionary change shown here, back towards the more virulent genotype. Given these considerations, we propose that the ORF1 recombination hotspots associated with different mutations observed in CPMMV have actually been selected for, as they would permit rapid adaptation to different hosts.

Data accessibility

All viral sequences obtained in this work were submitted to National Center for Biotechnology Information (NCBI) and the accessions numbers are available in Supplementary Table S8.

Author statement

L.G.Z., C.M.C., F.M.Z, S.L.E., A.J. and E.S.G.M. conceived the study; L.G.Z., T.A.T., M.S.A., R.R.P.N.L., D.F.M. gathered the data; L.G.Z, T.B.M., E.S.G.M., A.J., S.L.E and C.M.C analyzed the data; L.G.Z, C.M.C, F.M.Z. and S.L.E. wrote the manuscript with input and revisions from all coauthors.

Declaration of Competing Interest

The authors report no declarations of interest

Acknowledgements and funding information

We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) who supported this study. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2021.198389.

References

Alabi, O.J., Martin, R.R., Naidu, R.A., 2010. Sequence diversity, population genetics and potential recombination events in grapevine rupestris stem pitting-associated virus in Pacific North-West vineyards. J. Gen. Virol. 91, 265–276.
Alizon, S., Hurford, A., Mideo, N., Van Baalen, M., 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. J. Evol. Biol. 22, 245–259.
Almeida, A.M.R., Piuga, F.F., Kitajima, E.W., et al., 2003. Necrose da haste da soja. Série Documentos-Embrapa Soja 221, 1–48.
Anderson, R.M., May, R.M., 1982. Coevolution of hosts and parasites. Parasitology 85, 411–426.
Bardel, C., Danjean, V., Genin, E., 2006. ALTree: association detection and localization of susceptibility sites using haplotype phylogenetic trees. Bioinformatics 22, 1402–1403.
Bedhomme, S., Lafforgue, G., Elena, S.F., 2012. Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. Mol. Biol. Evol. 29, 1481–1492.
Belliure, B., Janssen, A., Maris, P.C., Peters, D., Sabelis, M.W., 2005. Herbivore arthropods benefit from vectoring plant viruses. Ecol. Lett. 8, 70–79.
Bernet, G.P., Elena, S.F., 2015. Distribution of mutational fitness effects and of epistasis in the 5' untranslated region of a plant RNA virus. BMC Evol. Biol. 15, 274.
Berngruber, T.W., Froissart, R., Choisy, M., Gandon, S., 2013. Evolution of virulence in emerging epidemics. PLoS Pathog. 9, e1003209.
Burch, C.L., Chao, L., 2000. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature 406, 625–628.
Carrasco, P., Daros, J.A., Agudelo-Romero, P., Elena, S.F., 2007. A real-time RT-PCR assay for quantifying the fitness of tobacco etch virus in competition experiments. J. Virol. Methods 139, 181–188.
Carvalho, S.L., Nagata, T., Junqueira, B.R., Zanardo, L.G., Paiva, A.C., Carvalho, C.M., 2017. Construction of a full-length infectious cDNA clone of Cowpea mild mottle virus. Virus Genes 53, 137–140.

Cuevas, J.M., Willemsen, A., Hillung, J., Zwart, M.P., Elena, S.F., 2016. Temporal dynamics of intrahost molecular evolution for a plant RNA virus. Mol. Biol. Evol. 32, 1132–1147.
De Barro, P.J., Liu, S.S., Boykin, L.M., Dinsdale, A.B., 2011. *Bemisia tabaci*: a statement of species status. Annu. Rev. Entomol. 56, 1–19.
Diveki, Z., Salanki, K., Balazs, E., 2004. The necrotic pathotype of the cucumber mosaic virus (CMV) ns strain is solely determined by amino acid 461 of the 1a protein. Mol. Plant-microbe Interact. 17, 837–845.
Domingo, E., Holland, J.J., 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. 51, 151–178.
Edgar, R., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BioMed Central Bioinf. 5, 113.
Elena, S.F., 2016. Evolutionary transitions during RNA virus experimental evolution. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 371, 1701.
Elena, S.F., Bedhomme, S., Carrasco, P., et al., 2011. The evolutionary genetics of emerging plant RNA viruses. Mol. Plant Microb. Interact 24, 287–293.
Escriu, F., Fraile, A., Garcia-Arenal, F., 2003. The evolution of virulence in a plant virus. Evolution 57, 755–765.
Froissart, R., Doumayrou, J., Vuillaume, F., Alizon, S., Michalakis, Y., 2010. The virulence-transmission trade-off in vector-borne plant viruses: a review of (non-) existing studies. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 365, 1907–1918.
García-Arenal, F., Zerbini, F.M., 2019. Life on the edge: geminiviruses at the interface between crops and wild plant hosts. Annu. Rev. Virol. 6, 411–433.
Grubaugh, N.D., Andersen, K.G., 2017. Experimental evolution to study virus emergence. Cell 169, 1–2.
Hajimorad, M.R., Eggenberger, A.L., Hill, J.H., 2005. Loss and gain of elicitor function of soybean mosaic virus G7 provoking Rsv1-mediated lethal systemic hypersensitive response maps to P3. J. Virol. 79, 1215–1222.
Hasiów-Jaroszewska, B., Borodynko, N., Jackowiak, P., Figlerowicz, M., Pospieszny, H., 2011. Single mutation converts mild pathotype of the *Pepino mosaic virus* into necrotic one. Virus Res. 159, 57–61.
Hirata, H., Lu, X., Yamaji, Y., et al., 2003. A single silent substitution in the genome of Apple stem grooving virus causes symptom attenuation. J. Gen. Virol. 84, 2579–2583.
Huang, A.S., Baltimore, D., 1970. Defective viral particles and viral disease processes. Nature 226 (5243), 325–327.
Inoue-Nagata, A.K., Kormelink, R., Nagata, T., Kitajima, E.W., Goldbach, R., Peters, D., 1997. Effects of temperature and host on the generation of tomato spotted wilt virus defective interfering RNAs. Phytopathology 87, 1168–1173.
Iwaki, M., Thongmeearkon, P., Prommin, M., Honda, Y., Hibi, J., 1982. Whitefly transmission and some properties of *Cowpea mild mottle virus* on soybean in Thailand. Plant Dis. 66, 265–268.
Jiu, M., Zhou, X.P., Tong, L., et al., 2007. Vector-virus mutualism accelerates population increase of an invasive whitefly. PLoS One 2, e182.
Jombart, T., 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405.
Khatabi, B., Wen, R.H., Hajimorad, M.R., 2013. Fitness penalty in susceptible host is associated with virulence of Soybean mosaic virus on Rsv1-genotype soybean: a consequence of perturbation of HC-Pro and not P3. Mol. Plant Pathol. 14, 885–897.
Koh, S.H., Li, H., Sivasithamparam, K., Admiraal, R., Jones, M.G.K., Wylie, S.J., 2017. Evolution of a wild-plant tobamovirus passaged through an exotic host: fixation of mutations and increased replication. Virus Evol. 3, vex001.
Lalic, J., Elena, S.F., 2015. The impact of high-order epistasis in the within-host fitness of a positive-sense plant RNA virus. J. Evol. Biol. 28, 2236–2247.
Lalic, J., Cuevas, J.M., Elena, S.F., 2011. Effect of host species on the distribution of mutational fitness effects for an RNA virus. PLoS Genet. 7, e1002378.
Leach, J.E., Vera Cruz, C.M., Bai, J., Leung, H., 2001. Pathogen fitness penalty as a predictor of durability of disease resistance genes. Annu. Rev. Phytopathol. 39, 187–224.
Lenth, R.V., 2016. Least-squares means: the r package lsmeans. J. Stat. Softw. 69.
Luan, J.B., Yao, D.M., Zhang, T., et al., 2013. Suppression of terpenoid synthesis in plants by a virus promotes its mutualism with vectors. Ecology Letter 16, 390–398.
Martelli, G.P., Adams, M.J., Kreuze, J.F., Dolja, V.V., 2007. Family *Flexiviridae*: a case study in virion and genome plasticity. Annu. Rev. Phytopathol. 45, 73–100.
Martin, D.P., 2009. Recombination detection and analysis using RDP3. Methods Mol. Biol. 537, 185–205.
Martin, D.P., Lemey, P., Lott, M., et al., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.
Marubayashi, J.M., Yuki, V.A., Wutke, E.B., 2010. Transmissão do *Cowpea mild mottle virus* pela mosca branca *Bemisia tabaci* biótipo B para plantas de feijão e soja. Summa Phytopathol. 36, 158–160.
Mauck, K.E., 2016. Variation in virus effects on host plant phenotypes and insect vector behavior: what can it teach us about virus evolution? Curr. Opin. Virol. 21, 114–123.
Mauck, K.E., De Moraes, C.M., Mescher, M.C., 2010. Deceptive chemical signals induced by a plant virus attract insect vectors to inferior hosts. Proc. Nat. Acad. Sci. U. S. A. 107, 3600–3605.
Mauck, K., Bosque-Pérez, N.A., Eigenbrode, S.D., De Moraes, C.M., Mescher, M.C., 2012. Transmission mechanisms shape pathogen effects on host–vector interactions: evidence from plant viruses. Funct. Ecol. 26, 1162–1175.
McKenzie, C.L., 2002. Effect of *Tomato mottle virus* (ToMoV) on Bemisia tabaci biotype B (Homoptera: *aleyrodidae*) oviposition and adult survivorship on healthy tomato. Fla. Entomol. 85, 367–368.
Menzel, W., Winter, S., Vetten, H., 2010. Complete nucleotide sequence of the type isolate of *Cowpea mild mottle virus* from Ghana. Arch. Virol. 155, 2069–2073.
Moles, M., Delatte, H., Farreyrol, K., Grisoni, M., 2007. Evidence that *Cymbidium mosaic virus* (CymMV) isolates divide into two subgroups based on nucleotide diversity of coat protein and replicase genes. Arch. Virol. 152, 705–715.
Naidu, R.A., Gowda, S., Satyanarayana, T., et al., 1998. Evidence that whitefly-transmitted *Cowpea mild mottle virus* belongs to the genus *Carlavirus*. Arch. Virol. 143, 769–780.
Nylander, J.A.A., 2004. MrModeltest v2. Program Distributed by the Author.
Péréfarres, F., Thébaud, G., Lefeuvre, P., et al., 2014. Frequency-dependent assistance as a way out of competitive exclusion between two strains of an emerging virus. Proc. Royal Soc. B: Biol. Sci. 281.
Pinheiro, J., Bates, D., DebRoy, S., D S, 3.1-128 RCTnLaNMEMRpv, 2016. Nlme: Linear and Nonlinear Mixed Effects Models. R Package Version 3.1-128. http://CRAN.R-project.org/package=nlme.
R Core Team (2014): //URL http://www.R-project.org/.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Rosario, K., Capobianco, H., Ng, T.F., Breitbart, M., Polston, J.E., 2014. RNA viral metagenome of whiteflies leads to the discovery and characterization of a whitefly-transmitted carlavirus in North America. PLoS One 9 (1), e86748.
Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.
Sidhu, J.S., Mann, R.S., Butter, N.S., 2009. Deleterious effects of Cotton leaf curl virus on longevity and fecundity of whitefly, Bemisia tabaci (Gennadius). J. Entomol. 6, 62–69.
Tribodet, M., Glais, L., Kerlan, C., Jacquot, E., 2005. Characterization of Potato virus Y (PVY) molecular determinants involved in the vein necrosis symptom induced by PVYN isolates in infected Nicotiana tabacum cv. Xanthi. J. Gen. Virol. 86, 2101–2105.
Walker, H.L., Pirone, T.P., 1972. Number of TMV particles required to infected locally or systemically susceptible tobacco cultivar. J. Gen. Virol. 17, 241–243.
Wan, G., Jiang, S., Wang, W., et al., 2015. Rice stripe virus counters reduced fecundity in its insect vector by modifying insect physiology, primary endosymbionts and feeding behavior. Sci. Rep. 5, 12527.
Zanardo, L.G., Silva, F.N., Bicalho, A.A.C., et al., 2014a. Molecular and biological characterization of *Cowpea mild mottle virus* isolates infecting soybean in Brazil and evidence of recombination. Plant Pathol. 63, 456–465.
Zanardo, L.G., Silva, F.N., Lima, A.T., et al., 2014b. Molecular variability of *Cowpea mild mottle virus* infecting soybean in Brazil. Arch. Virol. 159, 727–737.
Zhang, T., Luan, J.B., Qi, J.F., et al., 2012. Begomovirus-whitefly mutualism is achieved through repression of plant defences by a virus pathogenicity factor. Mol. Ecol. 21, 1294–1304.