

RESEARCH

Open Access



Poly(A) introduced upstream of the upstream pseudoknot domain of *Tobacco mosaic virus* led to sequence deletion after serial passaging in host plants

Song Guo¹ and Sek-Man Wong^{1,2,3*}

Abstract

TMV (24A + UPD) is a mutant virus that a 24 nt internal poly(A) tract was introduced downstream of the coat protein (CP) gene in TMV genome. TMV (24A + UPD) induced more severe necrosis in *Nicotiana benthamiana* and its RNA level accumulated at higher level than that of TMV, indicating that the introduced 24-nt poly(A) tract enhanced virus replication and virulence. In this study, TMV (24A + UPD) was serially passaged 10 times in *N. benthamiana* to analyze the predominant sequence changes of the introduced internal poly(A) tract and the upstream pseudoknot domain (UPD) in its genome. Results showed that the introduced 24 nt of the poly(A) tract was extended from first to seventh rounds of serial passaging, but shortened from the eighth passage, and reduced to only three adenines at the tenth passage. Also, the nucleotide sequences downstream of the introduced poly(A) tract were deleted gradually during the 10 cycles of passaging. There were mutant viruses generated with partial deletion of CP gene during serial passaging, indicating that extension of the introduced internal poly(A) tract also led to deletion of coding gene sequence of TMV (24A + UPD). These results provided valuable information for our understanding of the dynamics in viral sequence changes to reach a tipping point in its host plants in order to maintain a co-existence relationship during virus evolution. In addition, the introduction of an internal poly(A) tract may be applied to other viruses to study virus evolution and natural selection in host plants.

Keywords: Internal poly(A) tract, Upstream pseudoknot domain (UPD), Serial passaging, Sequence equilibrium, Virus evolution

Background

Tobacco mosaic virus (TMV) infects many plant species and has a worldwide distribution. Numerous TMV strains cause serious economic losses in tobacco, tomato, and other crop plants (Scholthof et al. 2011). TMV genome is a single-stranded positive-sense RNA encoding four major proteins, including a 126-kDa RNA-dependent RNA polymerase (RdRp) protein and its translational read-through product of 183 kDa, a 30-kDa cell-to-cell movement protein (MP) and a 17.5-kDa coat protein (CP). The 5'

untranslated region (UTR) of TMV genomic RNA is about 70 nucleotides (nt) long. It contains 12 CAA repeats as a translation enhancer (Abel et al. 1986; Gallie et al. 1987) and is essential for efficient virus replication (Takamatsu et al. 1991). TMV RdRp binds to the 5' UTR of genomic RNA to enable viral replication (Kawamura-Nagaya et al. 2014). The entire 3'-UTR of TMV is involved in forming an extended tertiary structure containing a tRNA-like structure (TLS) and an upstream pseudoknot domain (UPD). TMV 3'UTR is functionally similar to a poly(A) tail of messenger RNAs (mRNAs) in enhancing TMV RNA stability and translation efficiency during replication (Gallie et al. 1991). The TLS of TMV can be aminoacylated and binds to several tRNA-specific enzymes (Más and Beachy 1999). TMV-UPD is a quasi-continuous double-helical stalk

* Correspondence: dbswsm@nus.edu.sg

¹Department of Biological Sciences, National University of Singapore, Singapore 117543, Republic of Singapore

²Temasek Life Sciences Laboratory, 1 Research Link, Singapore 117604, Republic of Singapore

Full list of author information is available at the end of the article



comprising three consecutive pseudoknots (PKs) (van Belkum et al. 1985). It can functionally substitute the poly(A) tail of mRNAs in plants and animals (Gallie and Walbot 1990). Eukaryotic elongation factor 1A interacts with TMV-UPD and mutational analysis revealed that the UPD conformation and some conserved primary sequence elements are required for the interaction (Zeenko et al. 2002). Similar tandem arrangement of structural units, UPD and TLS, are found in most tobamoviruses (except those without UPD), varying in size and number of PKs of the UPD (Gulyaev et al. 1994). There is a short nucleotide sequence upstream of the PKs in the 3'UTR of TMV. This sequence has been validated to contain a stem-loop structure that regulates virus replication (Guo and Wong 2018).

Hibiscus latent Singapore virus (HLSV) and *Hibiscus latent Fort Pierce virus* (HLFPV) are two tobamoviruses that possess a unique feature of an internal poly(A) tract within their 3'UTRs in place of UPD in the genomes (Kamenova and Adkins 2004; Srinivasan et al. 2005; Yoshida et al. 2014). The internal poly(A) tract of HLSV ranges from 77 to 96 nt long upstream of the TLS. Shortening of the internal poly(A) tract resulted in reduced HLSV RNA accumulation (Niu et al. 2015).

The minimal length required for HLSV CP expression in plants is 24 nt, suggesting that the length of internal poly(A) tract is also crucial for HLSV replication (Niu et al. 2015). The UPD can functionally substitute a poly(A) tail of mRNAs in plant and animal cells (Gallie and Walbot 1990). However, when the UPD is replaced by a 43 nt internal poly(A) tract in TMV genome, replication of the mutant virus (termed as TMV-43A) is significantly reduced, as compared with TMV. In addition, milder mosaic symptoms and absence of necrosis observed in TMV-43A inoculated *N. benthamiana* plants are different from the lethal symptoms induced by TMV infection (Niu et al. 2015). Moreover, plant host defense responses are lower in TMV-43A infected plants as compared to those of TMV, which may be correlated to the milder mosaic symptoms induced by TMV-43A infection (Guo and Wong 2017). When the introduced internal poly(A) tract and the UPD are coupled in TMV genome, different lengths of the internal poly(A) tract introduced into TMV 3'UTR lead to structural variations that affect virus accumulation and symptom expression. The mutant virus TMV (24A + UPD) attained a higher viral RNA accumulation than TMV and induced lethal necrotic symptoms in *N. benthamiana* plants more rapidly (Guo et al. 2015).

Plant RNA viruses generate genetic diversity among species as RNA replication is error-prone which leads to mutations and diverse populations. Mutation, recombination, and reassortment are three major forces driving

the evolution of viruses, and they also generate diversity in viral genomes, providing variants to adapt to different environment. Minor changes in a viral genome can result in significant differences in its phenotypes (Roosinck 1997). In this work, we serially passaged TMV (24A + UPD) in *N. benthamiana* plants for 10 times using mechanical inoculation. We monitored the predominant sequence variations of progeny viruses in the non-coding region upstream of UPD and the length variations of the introduced internal poly(A) tract in each passaging. The introduced internal poly(A) tract in TMV (24A + UPD) gradually extended in length during serial passaging but reduced after the seventh passage. Three nucleotides were left in the introduced internal poly(A) tract at the tenth passage. For the non-coding region sequence, deletion was progressive after each serial passaging. Finally, a total of 23 nucleotides downstream of the introduced poly(A) tract were deleted at the tenth passage. Based on symptom comparison with that of TMV (24A + UPD), the passaged mutant virus became less virulent after 10 rounds of passaging. The predicted RNA secondary structure of the predominant mutant virus in the tenth passage is different from that of all other major mutant viruses. All these results indicated that the introduced poly(A) tract in TMV (24A + UPD) ushered virus evolution through serial passaging, leading to selection of a mutant virus that allows co-existence with its host plants.

Results

The introduced internal poly(A) tract of TMV (24A + UPD) was first extended and then shortened during serial passaging

An internal poly(A) tract of 24 nt long was introduced downstream of CP gene sequence in TMV genome and termed as TMV (24A + UPD) (Fig. 1a). During serial passaging of TMV (24A + UPD) in *N. benthamiana* plants (Fig. 1b), total RNA was isolated from each passaging of inoculated plants and used to determine the length of the internal poly(A) tract through RT-PCR and DNA sequencing. According to the primers targeted regions in TMV genome (Fig. 1c), the amplified PCR fragment included sequences encompassing partial MP gene, complete CP gene, the introduced internal poly(A) tract and the entire 3' UTR of TMV (24A + UPD).

In all the 10 cycles of passaging, infected plants showed leaf necrosis and stem collapse, leading to plant death within a few days. In the first and second rounds of passaging, the inoculated plants showed severe leaf necrosis at 5 days post inoculation (dpi) and their leaves collapsed (dead) within the next 2 days. Along with serial passaging, the appearance of necrotic symptoms in the inoculated plants was delayed, which led to plant death by 11 dpi. However, in the inoculated plants at the

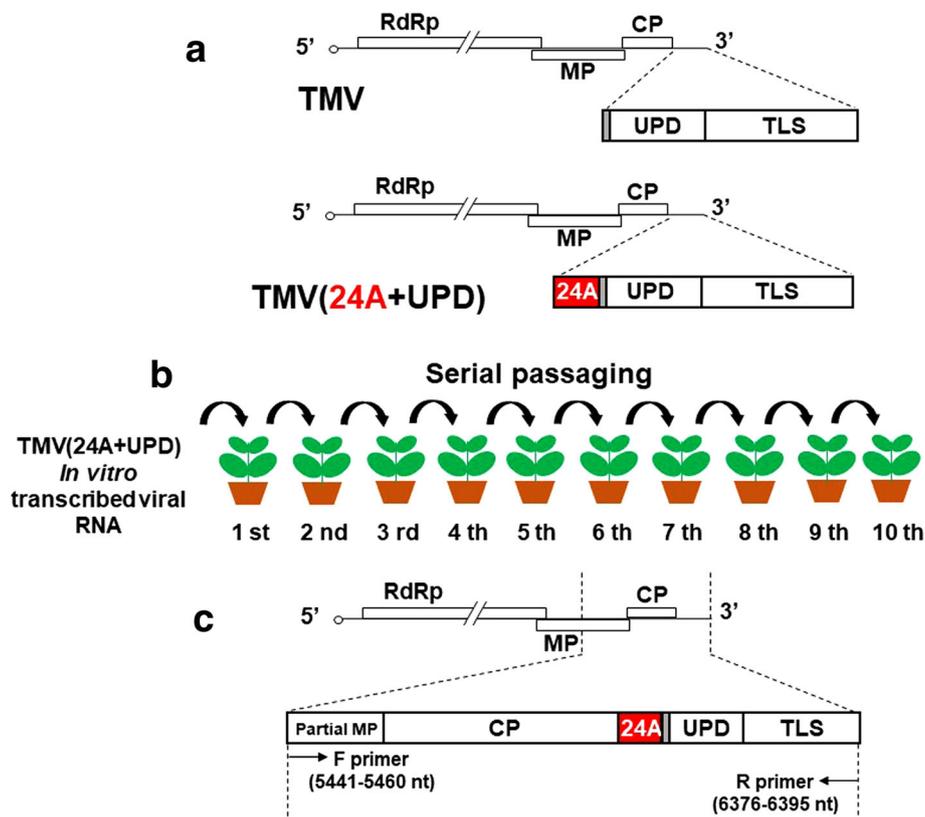


Fig. 1 **a** Schematic structures of TMV and TMV (24A + UPD) showing the internal poly(A) tract, upstream pseudoknot domain (UPD) and a short non-coding sequence (in grey color) between the internal poly(A) tract and the UPD. **b** Serial passaging of TMV (24A + UPD) in *N. benthamiana* plants. **c** Primers targeting regions in TMV (24A + UPD) genome are indicated

tenth-round of passaging, necrosis appeared on 5 dpi was milder than that of the first few cycles of passaging. One of the inoculated plants was still alive at 11 dpi (Fig. 2).

The viral CP was detected from total protein extracted from upper leaves of inoculated plants in each passaging (Fig. 3a), and the major PCR band of approximately 1 kb in size was purified and sequenced (Fig. 3b). DNA sequencing results of PCR products from each round of passaging showed that the introduced internal poly(A) tract extended from 24 nt to > 100 nt in the first few rounds of passaging but subsequently reduced rapidly (Fig. 4a). The exact lengths of internal poly(A) tract in clones from each round of passaging are listed in Additional file 1: Table S1.

The average length of internal poly(A) tract was calculated. The initially introduced 24 nt internal poly(A) tract in TMV (24A + UPD) was extended to ~ 40 nt (average length: 37 nt) after inoculated to *N. benthamiana* plants in the first round of passaging. Later, the introduced internal poly(A) length increased to ~ 100 nt (average length: 90 nt) until the seventh cycle of passaging and rapidly reduced to ~ 30 nt (average length: 32

nt) at the eighth cycle of passaging. Although the length of the introduced internal poly(A) tract again increased to ~ 55 nt (average length: 55 nt) at the ninth cycle of passaging, the internal poly(A) length showed only 3 nt left at the tenth cycle of passaging (Fig. 4b).

Nucleotides downstream of introduced internal poly(A) tract in TMV (24A + UPD) were deleted during serial passaging

The 24 nt internal poly(A) tract was introduced into the junction between the CP gene and a short non-coding sequence upstream of UPD (Fig. 5a). Therefore, it is reasonable to assume that the short non-coding sequence and the UPD may be affected by extension of the internal poly(A) tract during serial passaging. DNA sequencing results of clones from each round of passaging showed that a few nucleotides downstream of the introduced internal poly(A) tract were substituted by adenine (A) or uracil (U). The exact mutated nucleotides and their positions are listed in Additional file 1: Table S1. The mutation rate of the sequence downstream of poly(A) tract also increased with the extension of the introduced internal poly(A) tract during serial passaging

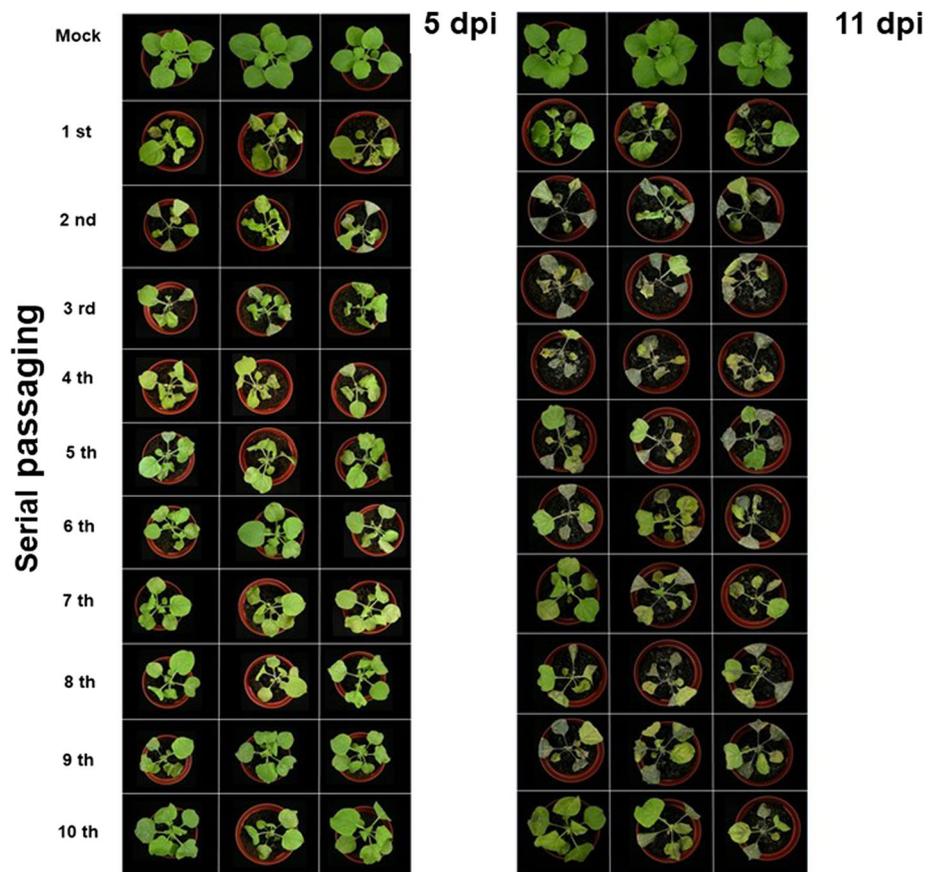


Fig. 2 Symptoms of TMV (24A + UPD) inoculated *N. benthamiana* plants during 10 rounds of serial passaging of TMV (24A + UPD) at 5 dpi and 11 dpi, respectively

(Fig. 5b). Among the 32 randomly selected clones from the first-round of inoculation, five clones changed in short non-coding sequences, with two nucleotides substituted by an A nucleotide. From the sixth serial passaging, nucleotides downstream of the introduced internal poly(A) tract in all 32 randomly selected clones were substituted by A or U nucleotide. Similar nucleotide substitutions were observed in subsequent serial passaging (seventh, eighth, ninth and tenth). The number of nucleotide mutation increased with serial

passaging. Of the five mutated clones of the first-round inoculation, four clones showed only one nucleotide of guanine (G) was substituted by an A nucleotide. However, in the second round of passaging, there were six clones with one G changed to A nucleotide. The number of a single nucleotide change increased in subsequent serial passaging. In the fifth serial passaging, four nucleotides downstream of the introduced internal poly(A) tract were deleted. In the eighth passaging, extensive deletion of 22, 25 and 42 nt were observed.

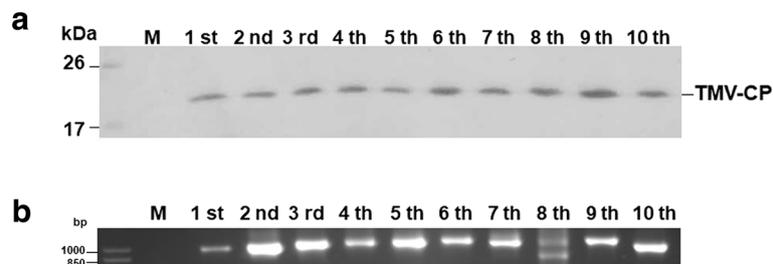


Fig. 3 a Detection of TMV (24A + UPD) coat protein (CP) in inoculated plants during 10 rounds of serial passaging. **b** Detection of PCR products of approximately 1 kb in size from 10 rounds of serial passaging of TMV (24A + UPD). M represents mock

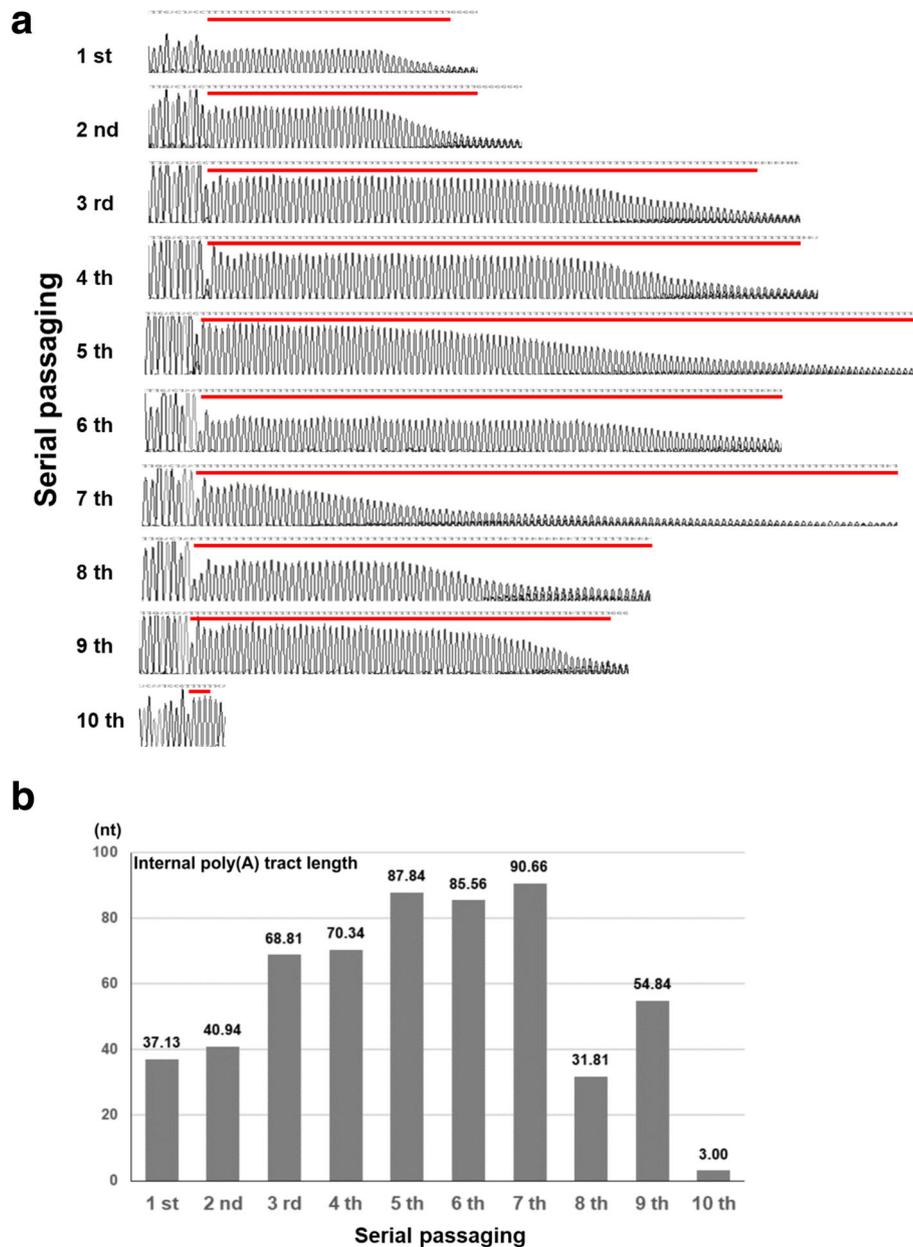


Fig. 4 a DNA sequencing results of 10 rounds of serial passaging of TMV (24A + UPD) in *N. benthamiana* plants, showing variations of internal poly(A) tract length (red line) in viral genome. **b** Average length of internal poly(A) tract in 10 rounds of serial passaging of TMV (24A + UPD)

A marked difference in sequence deletion was observed between the eighth and ninth passaging. While large sequence deletion was observed in several selected clones, all 32 clones of the ninth passaging showed only two nucleotides of G were substituted by A or U nucleotide. In the tenth-round of passaging, all 23 nt downstream of the introduced internal poly(A) tract were deleted from the 32 randomly selected clones (Table 1). The diversity of mutated viruses with different numbers of deleted nucleotides increased from the first to the eighth passaging. In the ninth and tenth passaging, all

clones showed the same number of deleted nucleotides, indicating an end to sequence deletion in TMV (24A + UPD) virus population after tenth serial passaging (Fig. 6).

Partial CP gene sequences were deleted with the introduced internal poly(A) tract during serial passaging
A 24 nt internal poly(A) tract was introduced downstream of CP sequence in TMV (24A + UPD) genome. Progeny viruses with deletion of the CP gene were first detected in the fifth round of passaging. The 3'end of

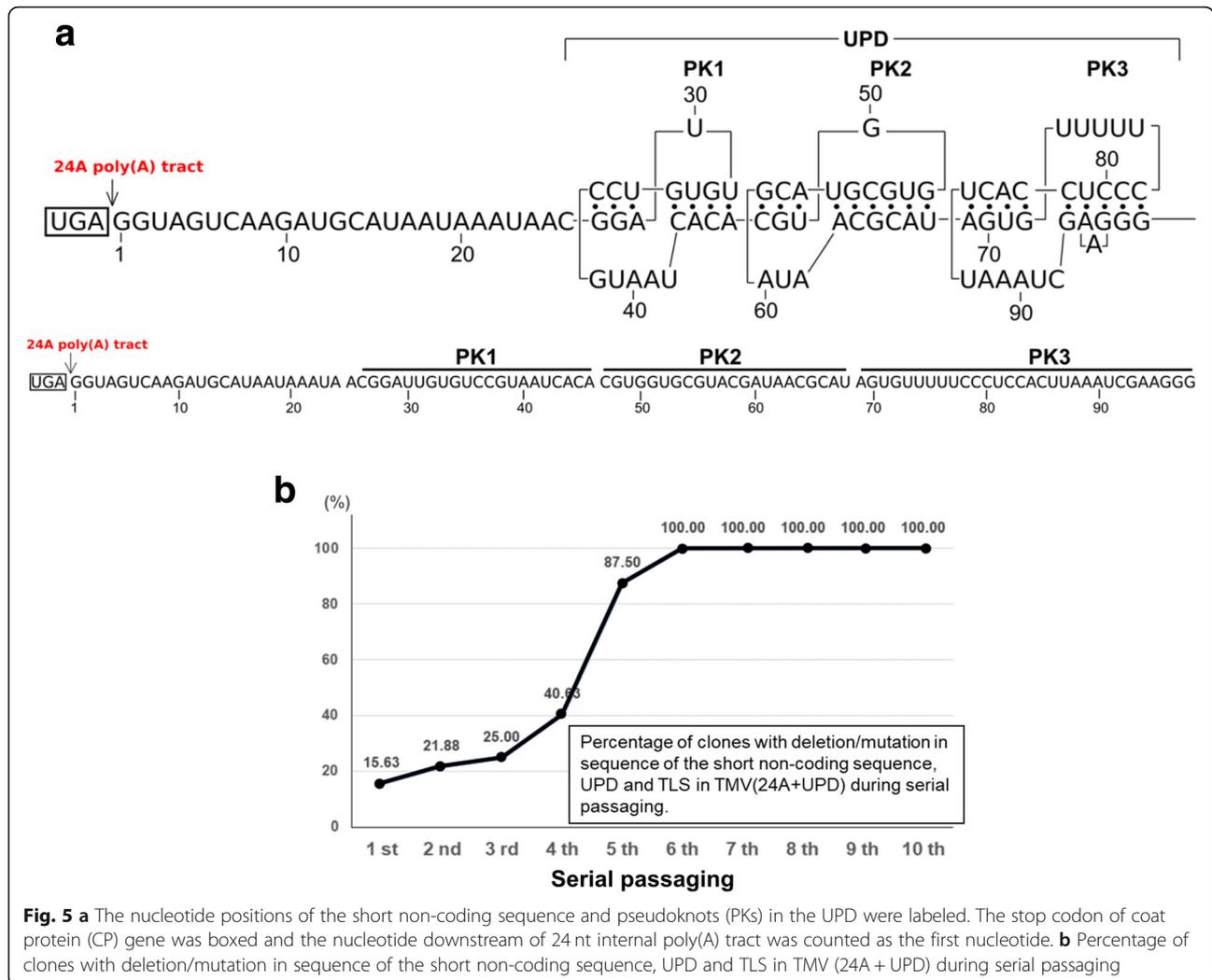


Table 1 Variation of clones with nucleotide deletion/mutation in sequence of the short non-coding sequence, UPD and TLS in TMV(24A+UPD) during serial passaging

Serial passaging	Number of clones with nucleotide deletion/mutation
1st	5 (1 nt: 4; 2 nt: 1)
2nd	7 (1 nt: 6; 2 nt: 1)
3rd	8 (1 nt: 7; 2 nt: 1)
4th	13 (1 nt: 8; 2 nt: 5)
5th	28 (1 nt: 12; 2 nt: 13; 3 nt: 2; 4 nt: 1)
6th	32 (2 nt: 28; 3 nt: 2; 4 nt: 1; 10 nt: 1)
7th	32 (1 nt: 2; 2 nt: 20; 3 nt: 8; 4 nt: 2)
8th	32 (2 nt: 20; 3 nt: 5; 22 nt: 2; 25 nt: 3; 42 nt: 2)
9th	32 (2 nt: 32)
10th	32 (23 nt: 32)

the CP gene was partially deleted. In the eighth round of passaging, most of the clones showed that the CP gene sequences were deleted from 79 nt to 293 nt (Additional file 1: Table S1). During the ninth and tenth serial passaging, CP gene deletion did not continue further.

Predicted RNA secondary structure of the predominant mutant virus in each round of passaging

Among the mutant viruses arising from each round of passaging, the deletion/mutation mainly occurred in the short non-coding sequence downstream of the introduced internal poly(A) tract and upstream of the UPD. The nucleotide sequence of the predominant mutant virus in each round of passaging was determined. The sequence included CP gene stop codon, the average length of internal poly(A) tract, and the predominant deletion/mutation sequence in the non-coding sequence from 32 clones randomly selected from each passaging (Fig. 7).

The predicted RNA secondary structure of the predominant mutant in each passage was analyzed. The

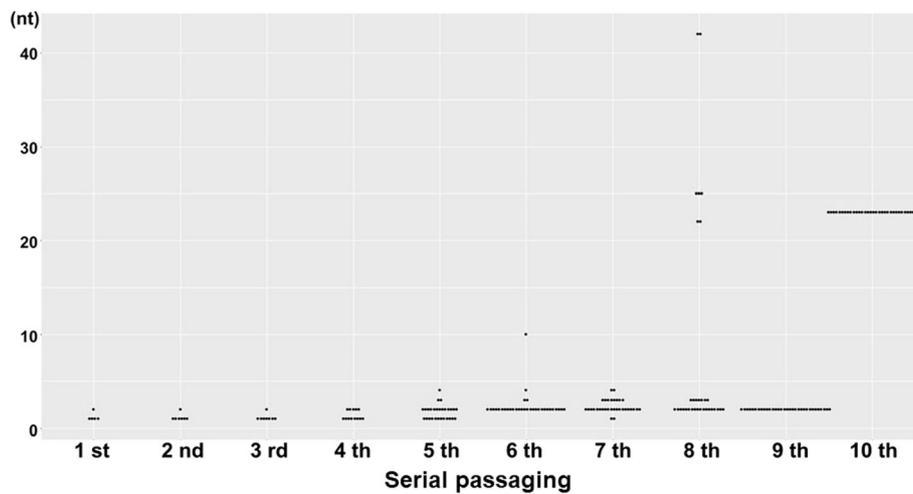


Fig. 6 Variation of deletion/mutation nucleotide numbers in the UPD region during 10 rounds of serial passaging of TMV (24A + UPD). The x-axis represents the number of serial passaging and the y-axis represents the deleted/mutated nucleotide numbers. The black dots in each round of serial passaging represent the clones with nucleotide deletion/mutation in the sequence downstream of internal poly(A) tract

sequence used for RNA secondary structure prediction included partial CP gene sequence, which forms a stem-loop structure with its 3' non-coding sequence (Guo and Wong 2018), the average length of internal poly(A) tract, the predominant deletion/mutation sequence in the non-coding sequence from the 32 clones in each round of passaging, UPD and TLS.

The predicted RNA secondary structure of the predominant virus mutant in the first and second rounds of passaging was similar to that of TMV (24A + UPD). The structure began to change at the third-round of passaging. From the third to the seventh round of passaging, the structures of the major mutant viruses were similar to each other. However, the structure

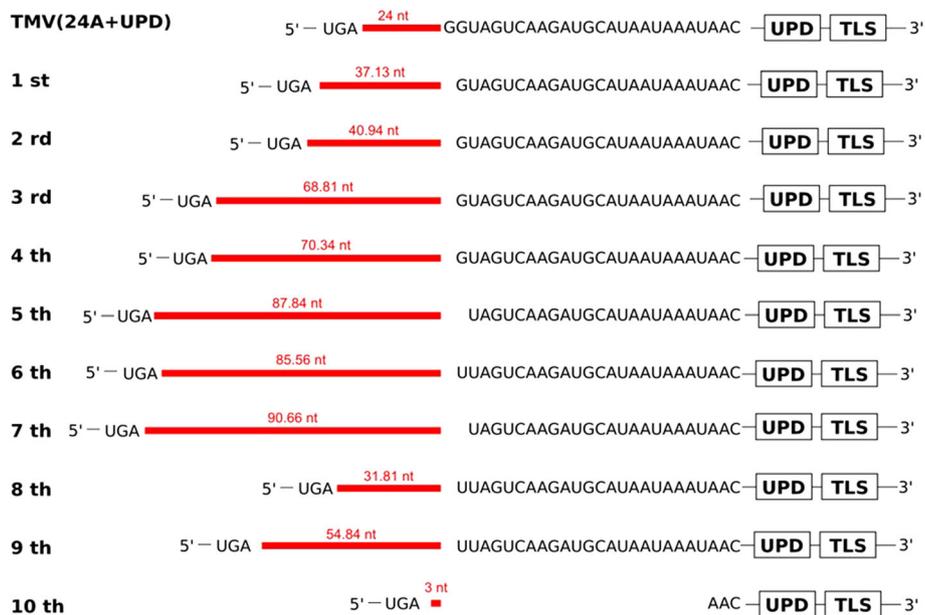


Fig. 7 The nucleotide sequence of the predominant mutant virus in each round passaging. The sequence included CP stop codon, the average length of poly(A) tract, the most occurred deletion/mutation sequence type in the non-coding sequence from the 32 clones in each round passaging, UPD and TLS

of the predominant mutant in the eighth and ninth rounds of passaging reverted to that of the first and second rounds of passaging. The predicted RNA structure of mutant virus in the tenth-round passaging was totally different from others (Additional file 2: Figure S1).

Discussion

Evolution of TMV mutant populations during serial passaging

TMV is the type member of the genus *Tobamovirus*. Studies on tobamovirus populations have showed that RNA sequence stability and rapid evolution are dependent on evolutionary conditions placed on the virus. Some studies reported that tobamoviruses are able to maintain its original sequence after passages and showed only a relatively small sequence diversity (Rodríguez-Cerezo and García-Arenal 1989; Rodríguez-Cerezo et al. 1991). However, others showed that tobamoviruses evolve rapidly from local lesions to create a great variety of variants (García-Arenal et al. 1984). Dramatic changes in the TMV population may occur in different host plants. Mixed TMV populations have occurred along with callus development and organogenesis in tissue cultures (Khan and Jones 1989a, 1989b). Genome evolution study of TMV populations during passaging in different host plants showed that host shift reduced TMV evolutionary rates (Kearney et al. 1999). Studies above indicated that when sufficient selective pressures were applied, TMV is capable of rapid evolution. In our study, it was showed that TMV (24A + UPD) is capable of rapid evolution, as the length changes of the introduced internal poly(A) tract in viral genome, mutations in the CP gene sequence and deletion of the non-coding sequences downstream of the introduced internal poly(A) tract were observed during serial passaging in *N. benthamiana* plants.

A previous report has shown that after long-term passaging of TMV in *N. tabacum* plants, less than 10% of total sequences examined were mutated. Most of the nucleotide changes do not affect amino acids sequence (Kearney et al. 1999), which indicates that TMV genomic sequence is relative stable during serial passaging. Since nucleotide deletions/mutations occurred in TMV (24A + UPD) during serial passaging were resulted from an introduced internal poly(A) tract that is not spontaneously occurred during TMV replication, wild-type TMV is not considered to be a proper control for this study (Roossinck, Elena and Wong, personal communication, 2019).

TMV and TMV (24A + UPD) require RNA-dependent RNA polymerase (RdRp) for replication. The intrinsically high error rate of RdRp contributes to the generation of population diversity that facilitates virus adaptation and evolution (Beaucourt et al. 2011). After introducing an

internal poly(A) tract into the TMV genome, TMV (24A + UPD) displayed a higher RNA accumulation than TMV. However, with increased length of introduced poly(A) tract, virus accumulation was reduced, and symptoms induced on plants became milder. TMV (62A + UPD) is unable to accumulate in *N. benthamiana* plants after inoculation (Guo et al. 2015). This indicates that the length of internal poly(A) tract could affect virus replication. Similar reports on *Tick-borne encephalitis virus* (TBEV) showed that a short poly(A) tract is genetically more stable, as compared to the virus with a long poly(A) tract (Asghar et al. 2016). During the serial passaging of TMV (24A + UPD) in *N. benthamiana* plants, the introduced internal poly(A) tract extended from 24 nt to ~100 nt. Utilization of RNA polymerase slippage for gene expression in positive-sense RNA virus has been reported among plant viruses belonging to *Potyviridae* (Olsper et al. 2015, 2016; Rodamilans et al. 2015; Untiveros et al. 2016). The slippage of viral RNA polymerase results in a population of transcripts with an additional A residue inserted in a highly conserved GAAAAA sequence, which enables expression of a viral protein that is essential for virus cell-to-cell movement (Olsper et al. 2015). TMV (24A + UPD) as a positive-sense RNA virus, the extension of its introduced internal poly(A) tract is believed to be due to slippage of viral RNA polymerases during virus replication.

Extension of internal poly(A) tract led to deletion of its downstream sequences during serial passaging

Many tobamoviruses showed that their UPD sequences are composed of three PKs. Virtually all the phylogenetically conserved primary sequences are present in the PKs of the UPD. The UPD is responsible for enhancing translational efficiency, similar to a poly(A) tail (Gallie and Walbot 1990). Both the conserved primary sequence within the UPD and its higher-ordered structure are necessary for the function of UPD in regulating translation (Leathers et al. 1993). The double-helical segment I, from 6257 nt to 6277 nt in the PK3 of UPD in TMV-L (a tomato strain) has been shown to be a structural feature essential for viral replication (Takamatsu et al. 1990). Structural analysis showed that the secondary structure of PKs in the UPD of TMV (24A + UPD) was similar to that of TMV, which indicated that the introduced 24 nt internal poly(A) tract did not alter the conformation of 3'UTR of viral genome (Guo et al. 2015). However, with additional serial passaging, mutations of the nucleotide sequences downstream of poly(A) tract also increased. Twenty-three nucleotides downstream of internal poly(A) tract were deleted near the PK1 of the UPD after the tenth passaging. Part of the short sequence between the introduced internal poly(A) tract and the UPD consists of a stem-loop structure in TMV genome to regulate viral

replication. Deletion of the stem-loop structure in TMV showed an enhanced virus infectivity and RNA accumulation (Guo and Wong 2018). However, at the tenth round of passaging, virulence of the final mutant virus and symptoms on plants were reduced, as compared to those of TMV (24A + UPD) infected plants in the first round of passaging. As the stem-loop structure consisted of partial CP gene and partial short non-coding sequences in TMV, the 24 nt introduced internal poly(A) tract downstream of the CP gene in TMV (24A + UPD) disrupted the formation of the stem-loop structure.

Since PKs are stable structures formed by base-pairing, mutations resulted from extension of internal poly(A) tract during virus passaging are limited. Some mutant viruses with deletion in the UPD were found in the eighth passage. Up to 42 nt downstream of the introduced internal poly(A) tract were deleted in those clones. Mutant viruses with a substantial sequence deletion in the UPD did not appear in subsequent passaging, indicating that further sequence deletion of important elements in the PKs of UPD is not permissible during virus evolution.

Changes of RNA secondary structure of mutant viruses during serial passaging

The predicted RNA secondary structures of major mutant viruses from serial passaging can be placed into three groups. The structures of major mutant viruses in the first, second, eighth and ninth rounds of passaging were similar to each other. The structures of major mutants from the third to the seventh rounds of passaging can be placed into one group. The structure of mutant virus in the tenth round of passaging was different from all others.

The RNA structure changes of mutants in each passage are caused by the introduced internal poly(A) tract. Previous study showed that introducing a 62 nt internal poly(A) tract into TMV genome could alter the RNA structure in its 3'UTR, resulting in a reduced viral RNA accumulation in *N. benthamiana* protoplasts and the inoculated plants remained symptomless (Guo et al. 2015). In this study, the lengths of poly(A) tract in the mutants from the first, second, eighth and ninth rounds of passaging were all shorter than 62 nt. Conversely, the lengths of poly(A) tract in the third to the seventh rounds of passaging were longer than 62 nt. The alterations of RNA structure in the virus mutants from each passaging were consistent with our previous study, in that longer internal poly(A) tract could alter RNA structure of 3'UTR of viral genome.

Among the major mutant viruses resulted from each round of passaging, one or two nucleotides were deleted/mutated in the short sequences between the internal poly(A) tract and UPD. The RNA structures of mutant viruses were not affected by the deletion/mutation.

However, in the tenth-round of passaging, 23 nt were deleted from the short sequence and the introduced internal poly(A) tract was reduced to three adenines. Therefore, the RNA structure of mutant virus in the tenth-round of passaging was different from that of all other mutants resulted from first to ninth rounds of serial passaging.

Generation of mutant viruses with deleted CP gene sequence during serial passaging

Defective interfering viruses arise through deletion, rearrangement, or recombination of a competent viral genome (Marriott and Dimmock 2010). Defective viruses compete with complete virus for replication and/or encapsidation factors and tend to attenuate viral symptoms. During TMV replication, the positive-strand RNA is used as template for negative-strand RNA synthesis and vice versa. Mutant viruses with CP gene deleted would affect virus replication and become defective interfering viruses. Their genomes lack a complete CP for virus assembly and long-distance movement in the host. Such mutant viruses are unable to move systemically in host plants and are less competitive. This virus population will subsequently be outcompeted by natural selection. By the eighth serial passaging, most of the clones showed truncated CP sequence and unable to move systematically. TMV requires functional CP for long-distance movement (Dawson et al. 1988; Saito et al. 1990; Hilf and Dawson 1993; Spitsin et al. 1999). TMV mutants with incomplete CP in the eighth round of passaging were not detected in the next round of passaging. It is believed that the mutant virus with a truncated CP is unable to move systemically to upper leaves.

Evolutionary question between the UPD and the internal poly(A) tract among tobamoviruses

Viral RdRps have a higher mutation rate than cellular polymerases (Duffy et al. 2008). Because of the lack of proofreading, the mutation rates of RdRps fall in the range of $\sim 2 \times 10^{-5}$ to 1×10^{-4} per nucleotide per replication event (Drake and Holland 1999). The high mutation rate leads to more defective genomes, but it allows the virus to evolve more rapidly through natural selection (Lauring et al. 2013; Paff et al. 2014). The RdRps of TMV (24A + UPD) replicate positive-strand RNAs into negative-strand RNAs which in turn produce multiple progeny-positive strands that are encapsidated into new virions. There were single nucleotide mutations in each serial passaging introduced by RdRps during virus replication. However, the effects of the introduced internal poly(A) tract to the sequence in TMV (24A + UPD) genome were predominately nucleotide deletions, rather than substitutions. We believe longer internal poly(A) tract forms RNA structure that blocks partial downstream non-coding sequences and the CP gene, leading

to RNA replication that skips sequences being blocked and generates progeny viruses with sequences deleted in the downstream non-coding region and the CP gene.

The introduced internal poly(A) tract has ushered the evolution of TMV (24A + UPD) viral genome to create different virus populations during serial passaging. TMV (24A + UPD) replication leads to different lengths of internal poly(A) tract, indicating generation of viral quasispecies. However, the virus populations started to reduce at the ninth serial passaging, and the introduced poly(A) sequence eventually reduced to three adenines upstream of a partially deleted UPD in the tenth serial passaging. Virus genetic bottlenecks and natural selection are contributing factors. Virus genetic bottlenecks are evolutionary events that reduce genetic variation of a population in a stochastic manner and result in small founding populations that lead to genetic drift (Li and Roossinck 2004). It takes place during systemic movement of TMV in leaves of the same plant (Sacristán et al. 2003).

Genetic drift reduces genetic variation in population, which potentially decreases the ability of a population to evolve in response to new selective pressures (Choi et al. 2001). The final result is elimination from the population of the fittest genotypes and accumulation of deleterious mutations. Therefore, it could lead to population extinction, such as shown experimentally in *Tobacco etch virus* (de la Iglesia and Elena 2007). Mutation accumulation and population extinction were also shown in reduced TMV population infecting *Nicotiana glauca*, caused by coinfection with *Tobacco mild green mosaic virus* (TMGMV) (Fraile et al. 1997).

Selection leads to differential reproduction of genetically distinct mutants or variants in a population and results in decreased population diversity. After selection, the fit variants will generate more progeny in the next generation than the less fit variants. Genetic drift and selection are not mutually exclusive and usually occur concurrently. In a large population, selection will push the frequency of a mutant upwards or downwards depending on whether the mutation is beneficial or harmful. In a small population, genetic drift will be dominant. It was reported that the number of TMV particles in a systemically infected tobacco leaf forms a large population (García-Arenal and McDonald 2003). Therefore, the number of TMV (24A + UPD) particles may also form a large population, as it replicates faster than TMV (Guo et al. 2015). Therefore, during serial passaging, selection appears to play a more important role in virus evolution than genetic drift. However, host plants also affect virus genomic combination, as viruses co-evolve with host plants. Certain mutations in viral genomes are only shown up in specific host plant species (Safari and Roossinck 2018). In our study, TMV (24A + UPD) was serially

passed in *N. benthamiana* plants and evolved into a new mutant virus. The symptoms developed in the tenth-round of passaging plants appeared milder than those of TMV (24A + UPD) infected plants in the first round of passaging. Co-evolution occurs between virus and specific host plants during serial passaging. Certain mutations and defective virus species are unable to pass on to the next round of passaging due to defects in virus systemic movement. In summary, this is the first study of virus evolution through introducing a foreign sequence [24 nt internal poly(A) tract] into viral genome to investigate changes of progeny viral sequences in host plants after multiple cycles of serial passaging.

Conclusions

Our results demonstrate that the introduced 24 nt internal poly(A) tract in TMV (24A + UPD) resulted in progeny sequences undergoing sequential mutations during serial passaging in *N. benthamiana* plants. After 10 rounds of passaging, a new mutant virus was generated which induced milder symptoms than TMV (24A + UPD). These results provided valuable information to our understanding of dynamic viral sequence changes in its host plants to select for progeny virus that allows it to co-exist with its hosts during virus evolution. Similar approach may be applied to other viruses to study virus evolution and natural selection in host plants.

Methods

Serial passaging of TMV (24A + UPD) in *N. benthamiana* plants

N. benthamiana plants were grown in a plant growth room at 24 °C with a photoperiod of 16 h light/8 h dark. The first round of inoculation was performed using in vitro transcribed RNA of TMV (24A + UPD) generated by mMessage T7 transcription kit (Ambion, USA). *KpnI* linearized TMV (24A + UPD) cDNA clone was used as template for in vitro transcription and 1 µg of transcribed RNA was used for inoculation with 1 × GKP buffer (50 mM glycine; 30 mM K₂HPO₄, pH 9.2; 1% bentonite; 1% celite) onto three *N. benthamiana* plants at the growth stage of 4 weeks old.

When necrotic symptoms first appeared on the upper leaves of inoculated plants, inoculum was prepared by grinding 1 g of the infected upper leaves in 10 mL of 0.1 M phosphate buffer, pH 7.0, using chilled mortar and pestle. After centrifuging at 8500×g for 10 min at 4 °C, supernatant was collected for the next round of inoculation (passaging). Three plants were inoculated for each round of serial passaging. When symptoms appeared, upper leaves from the three inoculated plants were collected and pooled for RNA isolation, reverse-transcription, PCR amplification, cDNA cloning and sequencing.

PCR amplification and DNA sequencing of fragment containing introduced internal poly(A) tract

Total RNAs were extracted from the pooled upper leaves by TRIzol® reagent. Reverse transcription was performed using the SuperScript III Reverse Transcriptase kit (Invitrogen, USA) with primer R (5'-TGGGCCCTACC GGGGTAA-3'). PCR was performed using KAPA Taq PCR kit (Kapa Biosystems, USA) with primers R and F (5'-ACGTGAGAGACGGAGGGCCC-3'). PCR products were purified by FavorPrep Gel/PCR purification kit (Favorgen Biotech Corp, USA). Based on the positions of F (5441–5460 nt) and R (6376–6395 nt) primers binding to viral genome, a major PCR amplified band of expected size of ~1 kb was observed. It was excised from 1% agarose gel after electrophoresis and purified. Other sizes of unspecific PCR products from each round of passaging were not considered. The ~1 kb purified PCR product was ligated into pGEM-T Easy vector (Promega, USA) using T4 DNA ligase (ThermoFisher Scientific, USA) and transformed into *E. coli* DH5α competent cells. Colony PCR was performed to screen for positive clones from each plate to confirm successful cloning. Finally, a total of 32 positive clones were randomly selected and cultured for plasmid isolation using AxyPrep Plasmid Miniprep kit (Biolab, USA). Sequencing PCR was performed using Big-Dye™ Terminator v3.1 Cycle Sequencing kit (ThermoFisher Scientific, USA) with primer R. The nucleotide sequences were analyzed with the DNAMAN package (Version 9, Lynnon BioSoft, Quebec, Canada). The exact lengths of internal poly(A) tract in clones from each round of passaging are listed in Additional file 1: Table S1.

Figure 5a was adopted from the study of Zeenko et al. (2002) and drawn using Inkscape (<https://inkscape.org/>). Variation of deletion/mutation nucleotide numbers in the UPD region during 10 rounds of serial passaging of TMV (24A + UPD) (Fig. 6) was generated using R packages (R version 3.5.0) (<https://www.r-project.org/>) (R Core Team, 2018).

Prediction of RNA secondary structure of major mutant virus in each round of passaging

RNA fold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) was used for secondary structure prediction of major mutant virus in each round of passaging. Minimum free energy (MFE) and partition function was chosen as fold algorithms when performing structure prediction. The sequence used for structure prediction in each round of passaging is listed in Additional file 2: Figure S1, which included partial CP sequence, the average length of poly(A) tract, the most frequently occurred deletion/mutation sequence in the non-coding sequence from the 32 clones in each round of passaging, including UPD and TLS sequences.

Additional files

Additional file 1: Table S1. Analysis of sequencing results of all cDNA clones in 10 rounds of serial passaging of TMV (24A + UPD) in *N. benthamiana* plants, showing the length of internal poly(A) tract, deletion/mutation of UPD and CP sequences in TMV (24A + UPD) viral genome. (XLSX 23 kb)

Additional file 2: Figure S1. Predicted RNA secondary structure of the predominant mutant virus in each round passaging. The sequence for structure prediction includes partial CP sequence, which is used to form a stem-loop structure with 3'-non-coding sequence in TMV genome (Guo and Wong 2018), the average length of poly(A) tract, the most occurred deletion/mutation sequence type in the non-coding sequence from the 32 clones in each round passaging, UPD and TLS. (DOCX 1310 kb)

Abbreviations

dpi: Days post inoculation; PCR: Polymerase chain reaction; PK: Pseudoknot; RT: Reverse transcription; TLS: tRNA-like structure; TMV: *Tobacco mosaic virus*; UPD: Upstream pseudoknot domain

Acknowledgements

We thank Wei Liang Sim for his assistance in sequencing of TMV (24A + UPD) mutant progenies during serial passaging. This work was supported by the Singapore Ministry of Education Tier 1 Academic Research grant R-154-000-B23-114, Suzhou Municipal Science & Technology Bureau grant SNG2018039 and National Natural Science Foundation of China grant 31872639.

Authors' contributions

SG performed experiments, analyzed the results and drafted the manuscript. SMW supervised the research and edited the manuscript. Both authors have read and approved the manuscript.

Funding

Singapore Ministry of Education Tier 1 Academic Research grant R-154-000-B23-114 through National University of Singapore, Suzhou Science & Technology Bureau grant SNG2018039 and National Natural Science Foundation of China grant 31872639.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Biological Sciences, National University of Singapore, Singapore 117543, Republic of Singapore. ²Temasek Life Sciences Laboratory, 1 Research Link, Singapore 117604, Republic of Singapore. ³National University of Singapore (Suzhou) Research Institute, Suzhou 215123, China.

Received: 12 March 2019 Accepted: 30 June 2019

Published online: 25 July 2019

References

- Abel PP, Nelson RS, De B, Hoffmann N, Rogers SG, Fraley RT, et al. Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*. 1986;232:738–44.
- Asghar N, Lee Y-P, Nilsson E, Lindqvist R, Melik W, Kröger A, et al. The role of the poly(A) tract in the replication and virulence of tick-borne encephalitis virus. *Sci Rep*. 2016;6:39265.
- Beaucourt S, Bordería AV, Coffey LL, Gnädig NF, Sanz-Ramos M, Beeharry Y, et al. Isolation of fidelity variants of RNA viruses and characterization of virus mutation frequency. *J Vis Exp*. 2011;52:e2953.

- Choi IR, Hall JS, Henry M, Zhang L, Hein GL, French R, et al. Contributions of genetic drift and negative selection on the evolution of three strains of wheat streak mosaic tritrovirus. *Arch Virol*. 2001;146:619–28.
- Dawson WO, Bubrick P, Grantham GL. Modifications of the tobacco mosaic virus coat protein gene affecting replication, movement, and symptomatology. *Phytopathology*. 1988;78:783–9.
- de la Iglesia F, Elena SF. Fitness declines in *Tobacco etch virus* upon serial bottleneck transfers. *J Virol*. 2007;81:4941–7.
- Drake JW, Holland JJ. Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A*. 1999;96:13910–3.
- Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet*. 2008;9:267–76.
- Fraile A, Escriu F, Aranda MA, Malpica JM, Gibbs AJ, García-Arenal F. A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J Virol*. 1997;71:8316–20.
- Gallie DR, Feder JN, Schimke RT, Walbot V. Functional analysis of the tobacco mosaic virus tRNA-like structure in cytoplasmic gene regulation. *Nucleic Acids Res*. 1991;19:5031–6.
- Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TM. The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts in vitro and in vivo. *Nucleic Acids Res*. 1987;15:3257–73.
- Gallie DR, Walbot V. RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. *Genes Dev*. 1990;4:1149–57.
- García-Arenal F, McDonald BA. An analysis of the durability of resistance to plant viruses. *Phytopathology*. 2003;93:941–52.
- García-Arenal F, Palukaitis P, Zaitlin M. Strains and mutants of tobacco mosaic virus are both found in virus derived from single-lesion-passaged inoculum. *Virology*. 1984;132:131–7.
- Gulyaev AP, van Batenburg E, Pleij CW. Similarities between the secondary structure of satellite tobacco mosaic virus and tobamovirus RNAs. *J Gen Virol*. 1994;75:2851–6.
- Guo S, Kierzek E, Chen G, Zhou Y-J, Wong S-M. TMV mutants with poly(A) tracts of different lengths demonstrate structural variations in 3'UTR affecting viral RNAs accumulation and symptom expression. *Sci Rep*. 2015;5:18412.
- Guo S, Wong S-M. Deep sequencing analysis reveals a TMV mutant with a poly(A) tract reduces host defense responses in *Nicotiana benthamiana*. *Virus Res*. 2017;239:126–35.
- Guo S, Wong S-M. Disruption of a stem-loop structure located upstream of pseudoknot domain in *Tobacco mosaic virus* enhanced its infectivity and viral RNA accumulation. *Virology*. 2018;519:170–6.
- Hilf ME, Dawson WO. The Tobamovirus capsid protein functions as a host-specific determinant of long-distance movement. *Virology*. 1993;193:106–14.
- Kamenova I, Adkins S. Transmission, in planta distribution, and management of *Hibiscus latent Fort Pierce virus*, a novel tobamovirus isolated from Florida hibiscus. *Plant Dis*. 2004;88:674–9.
- Kawamura-Nagaya K, Ishibashi K, Huang Y-P, Miyashita S, Ishikawa M. Replication protein of tobacco mosaic virus cotranslationally binds the 5' untranslated region of genomic RNA to enable viral replication. *Proc Natl Acad Sci U S A*. 2014;111:E1620–8.
- Kearney CM, Thomson MJ, Roland KE. Genome evolution of tobacco mosaic virus populations during long-term passaging in a diverse range of hosts. *Arch Virol*. 1999;144:1513–26.
- Khan IA, Jones GE. Accumulation of necrotic lesion inducing variants in TMV-infected plantlets derived from leaf disks of *Nicotiana sylvestris*. *Can J Bot*. 1989a;67:984–9.
- Khan IA, Jones GE. Selection for a specific tobacco mosaic virus variant during bolting of *Nicotiana sylvestris*. *Can J Bot*. 1989b;67:88–94.
- Lauring AS, Frydman J, Andino R. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol*. 2013;11:327–36.
- Leathers V, Tanguay R, Kobayashi M, Gallie DR. A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. *Mol Cell Biol*. 1993;13:5331–47.
- Li H, Roossinck MJ. Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J Virol*. 2004;78:10582–7.
- Marriott AC, Dimmock NJ. Defective interfering viruses and their potential as antiviral agents. *Rev Med Virol*. 2010;20:51–62.
- Más P, Beachy RN. Replication of tobacco mosaic virus on endoplasmic reticulum and role of the cytoskeleton and virus movement protein in intracellular distribution of viral RNA. *J Cell Biol*. 1999;147:945–58.
- Niu S, Cao S, Huang L-J, Tan KC-L, Wong S-M. The length of an internal poly(A) tract of hibiscus latent Singapore virus is crucial for its replication. *Virology*. 2015;474:52–64.
- Olsper A, Carr JP, Firth AE. Mutational analysis of the *Potyviriidae* transcriptional slippage site utilized for expression of the P3N-PIPO and P1N-PISPO proteins. *Nucleic Acids Res*. 2016;44:7618–29.
- Olsper A, Chung BY-W, Atkins JF, Carr JP, Firth AE. Transcriptional slippage in the positive-sense RNA virus family *Potyviriidae*. *EMBO Rep*. 2015;16:995–1004.
- Paff ML, Stolte SP, Bull JJ. Lethal mutagenesis failure may augment viral adaptation. *Mol Biol Evol*. 2014;31:96–105.
- R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018. URL <http://www.R-project.org/>
- Rodamilans B, Valli A, Mingot A, León DS, Baulcombe D, López-Moya JJ, et al. RNA polymerase slippage as a mechanism for the production of frameshift gene products in plant viruses of the *Potyviriidae* family. *J Virol*. 2015;89:6965–7.
- Rodríguez-Cerezo E, Elena SF, Moya A, García-Arenal F. High genetic stability in natural populations of the plant RNA virus tobacco mild green mosaic virus. *J Mol Evol*. 1991;32:328–32.
- Rodríguez-Cerezo E, García-Arenal F. Genetic heterogeneity of the RNA genome population of the plant virus U5-TMV. *Virology*. 1989;170:418–23.
- Roossinck MJ. Mechanisms of plant virus evolution. *Annu Rev Phytopathol*. 1997;35:191–209.
- Sacristán S, Malpica JM, Fraile A, García-Arenal F. Estimation of population bottlenecks during systemic movement of tobacco mosaic virus in tobacco plants. *J Virol*. 2003;77:9906–11.
- Safari M, Roossinck MJ. Co-evolution of a persistent plant virus and its pepper host. *Mol Plant-Microbe Interact*. 2018;31:766–76.
- Saito T, Yamanaka K, Okada Y. Long-distance movement and viral assembly of tobacco mosaic virus mutants. *Virology*. 1990;176:329–36.
- Scholthof KB, Adkins S, Czosnek H, Palukaitis P, Jacquot E, Hohn T, et al. Top 10 plant viruses in molecular plant pathology. *Mol Plant Pathol*. 2011;12:938–54.
- Spitsin S, Stepelwski K, Fleysh N, Belanger H, Mikheeva T, Shivprasad S, et al. Expression of alfalfa mosaic virus coat protein in tobacco mosaic virus (TMV) deficient in the production of its native coat protein supports long-distance movement of a chimeric TMV. *Proc Natl Acad Sci U S A*. 1999;96:2549–53.
- Srinivasan KG, Min BE, Ryu KH, Adkins S, Wong SM. Determination of complete nucleotide sequence of *Hibiscus latent Singapore virus*: evidence for the presence of an internal poly(A) tract. *Arch Virol*. 2005;150:153–66.
- Takamatsu N, Watanabe Y, Iwasaki T, Shiba T, Meshi T, Okada Y. Deletion analysis of the 5' untranslated leader sequence of tobacco mosaic virus RNA. *J Virol*. 1991;65:1619–22.
- Takamatsu N, Watanabe Y, Meshi T, Okada Y. Mutational analysis of the pseudoknot region in the 3' noncoding region of tobacco mosaic virus RNA. *J Virol*. 1990;64:3686–93.
- Untiveros M, Olsper A, Artola K, Firth AE, Kreuzer JF, Valkonen JPT. A novel sweet potato potyvirus open reading frame (ORF) is expressed via polymerase slippage and suppresses RNA silencing. *Mol Plant Pathol*. 2016;17:1111–23.
- van Belkum A, Abrahams JP, Pleij CW, Bosch L. Five pseudoknots are present at the 204 nucleotides long 3' noncoding region of tobacco mosaic virus RNA. *Nucleic Acids Res*. 1985;13:7673–86.
- Yoshida T, Kitazawa Y, Komatsu K, Neriya Y, Ishikawa K, Fujita N, et al. Complete nucleotide sequence and genome structure of a Japanese isolate of hibiscus latent Fort Pierce virus, a unique tobamovirus that contains an internal poly(A) region in its 3' end. *Arch Virol*. 2014;159:3161–5.
- Zeenko W, Ryabova LA, Spirin AS, Rothnie HM, Hess D, Browning KS, et al. Eukaryotic elongation factor 1A interacts with the upstream pseudoknot domain in the 3' untranslated region of tobacco mosaic virus RNA. *J Virol*. 2002;76:5678–91.