



Genetic Diversity and Evolution of Satellite RNAs Associated with the Bamboo Mosaic Virus

Ing-Nang Wang^{1,2}, Chung-Chi Hu³, Ching-Wei Lee³, Sih-Min Yen³, Wen-Bing Yeh⁴, Yau-Heiu Hsu³, Na-Sheng Lin^{1*}

1 Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan, Republic of China, **2** Department of Biological Sciences, University at Albany, Albany, New York, United States of America, **3** Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan, Republic of China, **4** Department of Entomology, National Chung Hsin University, Taichung, Taiwan, Republic of China

Abstract

Satellite RNAs (satRNAs) are subviral agents that depend on cognate helper viruses for genome replication and encapsidation. Their negative impacts on helper viruses have been exploited to control plant viral diseases. SatBaMV is a commonly found satRNA associated with *Bamboo mosaic virus* (BaMV) that infects diverse bamboo species in the field. To investigate the genetic diversity and evolution of satRNAs, we examined seven satBaMV populations derived from five bamboo species and cultivars from Taiwan, China, and India and one from the greenhouse. We found 3 distinct clades among the seven populations. Clade I is consisted of all satBaMV isolates, except for those from *Dendrocalamus latiflorus* in Taiwan and *Bambusa vulgaris* in India, which belong to Clades II and III, respectively. Interestingly, nucleotide diversity was lower for Clade I than II and III. However, the nucleotide diversity did not seem to depend on bamboo species or geographic location. Our population genetic analyses revealed the presence of excessive low-frequency polymorphic sites, which suggests that the satBaMV population was under purifying selection and/or population expansion. Further analysis of *P20*, the only satBaMV gene that encodes a non-structural protein involved in the long-distance movement of satBaMV, showed evidence of purifying selection. Taken together, our results suggest that purifying selection against defective P20 protein is responsible at least in part for the evolution of the satBaMV genome.

Citation: Wang I-N, Hu C-C, Lee C-W, Yen S-M, Yeh W-B, et al. (2014) Genetic Diversity and Evolution of Satellite RNAs Associated with the Bamboo Mosaic Virus. PLoS ONE 9(10): e108015. doi:10.1371/journal.pone.0108015

Editor: A.L.N. Rao, University of California, Riverside, United States of America

Received: April 18, 2014; **Accepted:** August 20, 2014; **Published:** October 2, 2014

Copyright: © 2014 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by grants from the Institute of Plant and Microbial Biology (IPMB), Academia Sinica, Taiwan, to NSL. INW received funding from the National Science Council, Taiwan (NSC 101-2811-B-001-024), IPMB, and Academia Sinica during his sabbatical visit. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: nslin@sinica.edu.tw

Introduction

One of the most striking and under-appreciated differences between plant and animal viruses is the predominant association of satellite RNAs (satRNAs) with the plant viruses [1–3]. The genomes of satRNAs are small, usually <1,500 nucleotides (nt) [1]. Their sequences do not reveal any recognizable open reading frame (ORF) or show only a single ORF encoding a small protein. Because of their small genome size, satRNAs do not encode their own RNA-dependent RNA polymerases (RdRp) or coat proteins (CPs), so they depend on their cognate helper viruses for genome replication and encapsidation [1–5]. Therefore, satRNAs are commonly seen as molecular parasites [3,6] engaging in a zero-sum game by exploiting vital functions provided by their helper viruses.

One manifestation of this antagonistic interaction between satRNAs and their helper viruses is the attenuation of disease symptoms, presumably as a result of reduced viral titers in the host plants. Engineered satRNAs exploit this interfering capacity as antiviral agents for crop protection [7–10]. However, empirical studies have shown that the effects of satRNAs on their helper viruses can run the full gamut of all possible interactions, from antagonistic to beneficial [2,11], resulting in a wide range of disease symptoms [12]. One contributing factor to such variable

interactions may be the fascinating interplay among three parties: the satRNA, the helper virus, and the host defense mechanisms, most notably the RNA silencing pathway [1].

Besides the satRNAs associated with *Cucumber mosaic virus* (CMV) and *Turnip crinkle virus* [13], satRNAs associated with the *Bamboo mosaic virus* (satBaMV) are another of the intensively studied model systems. BaMV was first discovered in Brazil [14]. However, the most extensively studied isolates were first isolated from infected green bamboos (*Bambusa oldhamii* Munro) in Taiwan [15]. BaMV is a plant potexvirus with a filamentous, flexuous morphology of approximately 490 by 15 nm. The BaMV-O genome is a positive-stranded RNA of 6366 nt, with 5 ORFs [16]. The genome for satBaMV is 836 nt (without the poly-A tail) and encodes a single gene product of 183 amino acid residues [17] with similarity to the CP encoded by the satellite virus of *Panicum mosaic virus* [18]. The encoded 20-kDa protein is not required for satBaMV replication [19], but it does preferentially bind to the satBaMV RNA [20], presumably participating in the regulation of systemic movement of satBaMV in the host plant [19,21,22].

Because the replication of satBaMV depends on the RdRp encoded by BaMV, both the 5' [23,24] and 3' [25] untranslated regions (UTRs) of the satBaMV and BaMV RNAs, not

surprisingly, share structural similarities. Therefore, competition for the limited intracellular resources (*e.g.*, RdRp) can lead to an antagonistic interaction between these two entities, thus resulting in reduced BaMV titers and consequently attenuated symptoms [26–28].

However, not all satBaMV isolates behave antagonistically toward their helper virus. For example, the satBaMV isolate BSL6, first isolated from Ma bamboo (*Dendrocalamus latiflorus* Munro) [29], strongly interferes with BaMV replication [30], whereas BSF4, first isolated from *B. vulgaris* [17], does not seem to have a significant effect on BaMV [19,30]. Detailed genetic studies showed that various mutations at the 5′ apical hairpin stem loop (AHSL) region affect the ability of satBaMV to interfere with the replication of its helper virus [27,28]. A single nucleotide change of U to C at position 82 within the AHSL region of BSF4 can turn the non-interfering BSF4 into an interfering satRNA [28]. Therefore, the interfering phenotype is quite malleable and, presumably, is easily accessible via mutation.

In contrast to many detailed molecular studies, evolutionary and population studies of satRNAs are not as broad or as in-depth. The best-studied example is the various satRNAs associated with CMV [31–34]. At present, no systematic population study has investigated satBaMV, except for the previous surveys of field samples of bamboo species in Taiwan [23,29].

In this study, we collected, sequenced, and analyzed by phylogenetic and population genetics analyses of satBaMV isolates from five bamboo species and cultivars in three main geographic locations in Asia. We identified three distinct satBaMV clades and inferred the evolutionary forces responsible for the observed variations in our collection of satBaMV sequences.

Materials and Methods

Sample collection and preparation

Leaf samples showing visible mosaic symptoms from various bamboo species were collected from locations in northern (*B. vulgaris* Schrad. ex J. C. Wendl. and *D. latiflorus* Munro at the Taipei Botanical Garden; GPS coordinates: 24.91571, 121.67394), central (*B. oldhamii* Munro at Chung Hsing Lake of National Chung Hsing University at Taichung; GPS coordinates: 24.23321, 120.94174) and southern (*D. latiflorus* Munro cv. Mei-nung W. C. Lin at Pingtung; GPS coordinates: 22.67611, 120.49417) Taiwan in 2006; Hainan Island of China (*B. ventricosa* McClure; GPS coordinates: 19.00000, 109.50000) in 2007; and Delhi, India (*B. vulgaris* Schrad. ex J. C. Wendl.; GPS coordinates: 28.63531, 77.22496) in 2009. The presence of BaMV and satBaMV was checked by dot or northern blot analysis with BaMV and satBaMV-specific probes, respectively [19,35]. To ensure a fair representation of the satBaMV population, leaves from several individual bamboo plants in a given location were combined for total RNA extraction (see below). Figure 1 shows the locations of leaf samples and name designations for satBaMV populations.

To establish a baseline evolution in the greenhouse condition, *Nicotiana benthamiana* leaves were co-infected with 0.1 μg BaMV-S transcripts [36] and 0.1 μg BSF4 satBaMV RNA transcripts [19] obtained by T7 RNA polymerase (Promega) transcription, then harvested at 7 days post-infection (dpi).

RNA isolation, RT-PCR cloning, and sequencing

Total RNA was extracted from leaves as described [37] and was used as a template for RT-PCR reaction. The polyadenylation primer Tn836 (5′-CGTACCGAATTCT₁₅-3′, *Eco*RI site underlined) was used for first-strand cDNA synthesis with SuperScriptIII one-step RT-PCR system (Invitrogen), followed by PCR ampli-

fication with the high-fidelity Platinum *Taq* DNA polymerase (Invitrogen) and primer pairs for BS-5′ (5′-GAAAACCTCACCG-3′) and Tn836. The PCR product was then *Eco*RI-digested, gel-purified, and cloned into the *Stu*I/*Eco*RI sites of the pCass2 vector [38]. Plasmids were purified and used as templates for DNA sequencing to determine the identity of satBaMV isolates.

Sequence analyses

Clustal Omega [39] from EMBL-EBI (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) was used to align satBaMV sequences, with manual adjustment. The sequence alignment file is available upon request. Several recombination detection algorithms, such as RDP [40], GENECONV [41], Bootscan [42], MaxChi [43], Chimaera [44], SiScan [45], and 3Seq [46], all implemented in RDP4 beta 4.24 [47], were used to detect recombination among satBaMV sequences. MEGA 5.2.1 [48] was used to estimate nucleotide diversity and conduct selection tests. To determine the standard error associated with the estimated nucleotide diversity, 500 bootstraps were conducted. The online service, Intrapop neutrality tests (<http://www.wabi.snv.jussieu.fr/achaz/neutralitytest.html>), was used for Tajima's *D* [49] and Fu and Li's *D*₂* and *F** [50] tests. MrBayes 3.2 [51] was used to reconstruct phylogenetic relationships. Instructions in the command block was generated with the web service by R. Page and S. R. Santos (https://131.204.120.103/srsantos/mrbayes_form/index.html) and modified to fit our dataset. GTR (General Time Reversible) was used for the nucleotide substitution model and gamma-distribution the variable rate model. In this study, 1,000,000 generations were generated and 25% burn-in was used for each run. The resulting consensus trees were visualized by FigTree 1.3.1 [52].

Results

We collected seven leaf samples – six from various bamboo species from locations in Taiwan, China, and India and one from inoculated *Nicotiana benthamiana*, a commonly used tobacco plant in plant virus study, in the greenhouse (Figure 1) – to survey the population structures of satBaMVs. The satBaMV genomic sequences were determined by sequencing cloned RT-PCR products. A total of 728 cDNA clones were sequenced, with about 60 to 120 clones from each sample, although not all clones contained the satBaMV sequences (Table 1). Unexpectedly, more clones from the field samples did not contain the satBaMV sequences (13–34%) as compared with the greenhouse sample (~4%) (Table 1). After removing failed clones, 595 satBaMV-containing sequences were used for further analyses.

Baseline error rate from RT-PCR

To establish the baseline error rate introduced by RT-PCR, we used a plasmid-borne satBaMV cDNA as the template to generate RNA transcripts with the T7 RNA polymerase, followed by reverse transcription and PCR amplification under the same conditions used for field samples. We sequenced 18 clones and found only 4 substitution polymorphic sites, all singletons, in 17 full-length sequences of 836 nt and one length polymorphism of 732 nt (with deletion from nucleotides 189 to 292, based on the BSF4 sequence in GenBank accession no. AY205227). The mean nucleotide diversity [53], a population genetic statistic estimating the average number of nucleotide differences per site between two random sequences in a given population, was estimated to be 0.0006 ± 0.0003 .

To detect the presence of recombination in our data, we subjected aligned full-length satBaMV sequences to various recombination-detection algorithms (see Materials and Methods).

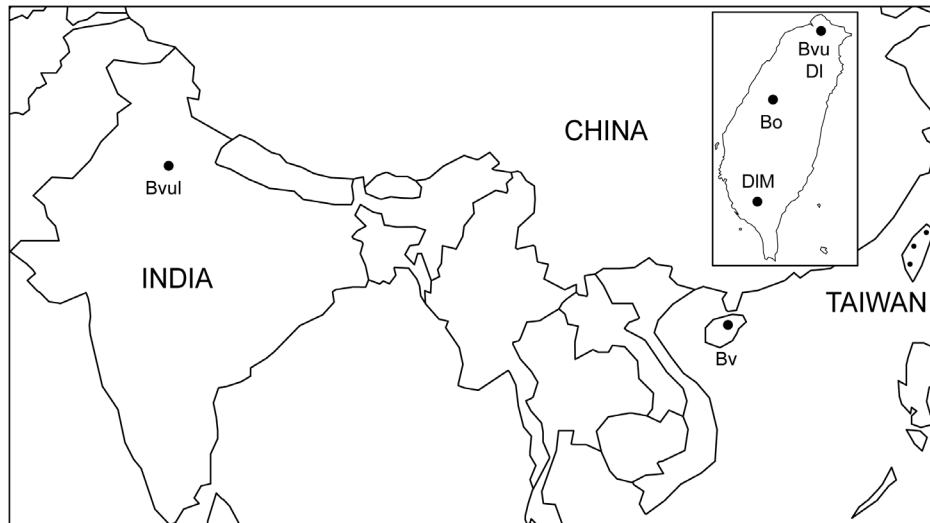


Figure 1. Map showing the geographic locations of collected bamboo samples. Inset shows an enlarged map of Taiwan for clarity. The name designations for each satBaMV populations are: Bvul, from *B. vulgaris* Schrad. ex J. C. Wendl. at Delhi, India; Bv, from *B. ventricosa* McClure at Hainan Island, China; Bvu, from *Bambusa vulgaris* Schrad. ex J. C. Wendl. at Taipei; DI, from *Dendrocalamus latiflorus* Munro at Taipei; Bo, from *B. oldhamii* Munro at Taichung, Taiwan; and DIM, from *D. latiflorus* Munro cv. Mei-nung W. C. Lin at Pingtung, Taiwan.
doi:10.1371/journal.pone.0108015.g001

Except for 2 sequences (L-89 and L-104), both from *D. latiflorus* Munro at the Taipei Botanical Garden, we found no evidence of recombination in our samples.

Overall, our results confirmed that RT-PCR introduces a small amount of sequence diversity in a given population study.

Apparent genome length polymorphism

Of the 595 sequences we analyzed, 496 more or less contained the full-length satBaMV sequences, with genome sizes ranging from 799 to 842 nt. Most of the genomes (69%) in the “full-length” category were 836 nt. The remaining 99 sequences all have deletions to various degrees and were classified into 7 categories based on genome length (Table 1). The group of 693–733 nt (“Short 1” category) was the second most common genome size found in all populations, except for populations from *D. latiflorus* Munro at the Taipei Botanical Garden (DI) and *B. vulgaris* from India (BvuI). Both populations had ~1% genomes in the “Short 1” category, whereas others had 7% to 22% sequences belonging to this category. SatBaMVs isolated from both *D. latiflorus* Munro (DI and DIM) hosts had a wider range of genome sizes, down to approximately 173–221 nt (“Short 7” category in Table 1).

As shown above, the 732-nt variant is likely a laboratory RT-PCR artifact. For the sequences in the 731–733-nt range, all had exactly the same deletion. Furthermore, individuals from a shorter category (e.g., “Short 3”) may be derived from a longer category (e.g., “Short 1”) (see Figure S1). Because of this uncertainty, we removed all deletion mutants for subsequent analysis.

Levels of genetic diversity among satBaMV populations

We adopted nucleotide diversity π [53] to measure the genetic diversity of satBaMV populations. The mean estimated nucleotide diversity ranged from 0.0014 ± 0.0006 (the greenhouse-inoculated *N. benthamiana*) to 0.0187 ± 0.0030 (*D. latiflorus* Munro, DI), approximately a 13-fold difference (Figure 2). Among the field samples, the lowest nucleotide diversity was with *B. vulgaris* Schrader ex Wendland (Bvu), with a value of 0.0020 ± 0.0006 .

To test whether these estimated values significantly differed, we used the Tukey-Kramer method for unplanned multiple comparisons [54] (Box 9.7, Part IV). SatBaMV populations from *D. latiflorus* Munro (DI) and *B. vulgaris* Schrader ex Wendland in India (Bvul) had the highest nucleotide diversity, and the rest had a lower value (Figure 2). Among our leaf samples were several contrasting satBaMV populations, derived from the same bamboo species but different locations or from the same location but different species. For example, for *B. vulgaris* Schrader ex Wendland, we had one sample from the Taipei Botanical Garden and one from India (Bvu vs. BvuI). As well, for *D. latiflorus* Munro, we had one sample from the Taipei Botanical Garden and one from Pingtung (DI vs. DIM). We also had satBaMV populations from *B. vulgaris* or *D. latiflorus* at the Taipei Botanical Garden (Bvu vs. DI). Interestingly, all pairwise comparisons revealed significant differences in nucleotide diversity (Figure 2).

Our analysis showed that nucleotide diversity among satBaMV populations varied greatly, indicating the presence of different demographic and/or evolutionary histories among these satBaMV populations. Furthermore, neither host species nor geographic location predominantly determined the level of nucleotide diversity within a given satBaMV population.

Phylogeny of the satBaMVs

The mean within-population nucleotide diversity was estimated at 0.0080 ± 0.0012 (or 0.0071 ± 0.0009 with the greenhouse sample), and the mean between-population diversity 0.0430 ± 0.0066 (or 0.0391 ± 0.0056 with the greenhouse sample), approximately a five-fold difference. This result, together with previous findings [23], suggests that each satBaMV population evolves independently, thus accumulating sequence divergence across all populations. This interpretation is further corroborated by the observation that most satBaMV sequences clustered with members from the same population, forming three well-supported clades (Figure 3). Clade I is consisted of all satBaMV populations, except most of those isolated from *D. latiflorus* Munro at the Taipei Botanical Garden (DI) and all of those from *B. vulgaris*

Table 1. Cloning effort and genome length distribution of satBaMV.

Population*	Bo	Bv	Bvu	DI	DIM	Bvul	B+F4
cDNA cloning effort							
Success	69	102	87	77	52	99	109
Fail	30	16	19	40	9	15	4
Fail rate	0.303	0.136	0.179	0.342	0.148	0.132	0.035
Percentage of genome length							
Genome size [†]	78.3	84.3	90.8	66.2	67.3	98.0	86.2
Short 1	21.7	6.9	9.2	1.3	15.4	1.0	13.8
Short 2				10.4	1.9	1.0	
Short 3		1.0			1.9		
Short 4				5.2			
Short 5		7.8		9.1	13.5		
Short 6				3.9			
Short 7				3.9			

*For population designation, see Figure 1 legend.

†Genome size in nucleotides. Full-length: 799–842; Short 1: 693–733; Short 2: 612–661; Short 3: 556; Short 4: 473–478; Short 5: 397–408; Short 6: 324–357; Short 7: 173–221. doi:10.1371/journal.pone.0108015.t001

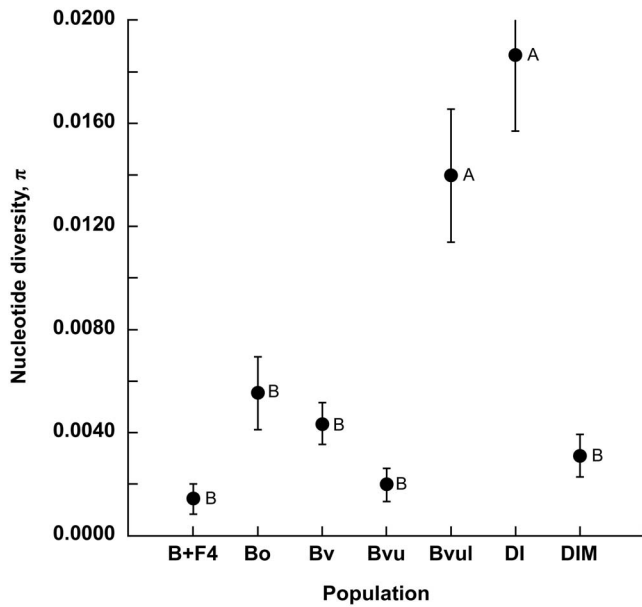


Figure 2. Nucleotide diversity of satBaMV populations. MEGA 5.2.1 [48] was used to estimate the nucleotide diversities and associated standard errors (vertical bars) of various satBaMV populations. Letters denote statistically different Tukey groups based on unplanned multiple comparisons with the Tukey-Kramer method [54]. For population designation, see Figure 1. B+F4 denotes the population from greenhouse-inoculated *N. benthamiana*. doi:10.1371/journal.pone.0108015.g002

Schrader ex Wendland in India (BvuI). Clade II is consisted of sequences found exclusively in *D. latiflorus* Munro (DI) at the Taipei Botanical Garden and Clade III sequences exclusively from *B. vulgaris* Schrader ex Wendland in India (BvuI). These three clades also corresponded with the levels of nucleotide diversity shown above, with all populations in Clade I having lower nucleotide diversity and both Clades II and III having significantly higher nucleotide diversity.

Most individual sequences tended to cluster with individuals from the same population (Figure 3), with some exceptions. The most notable exception was the well-supported clustering of sequences from *D. latiflorus* Munro cv. “Mei-nung” (DIM) in southern Taiwan and those from *B. ventricosa* McClure (Bv) in Hainan Island, China (Figure 3). This clustering suggests a recent migration event(s) of satBaMV between these two geographic locations. The second exception is the clustering of a single sequence from *B. vulgaris* (isolate a-34) and from *D. latiflorus* (isolate L-33), both from the Taipei Botanical Garden, with members from the greenhouse *N. benthamiana*. This second exception may not be too surprising, because the BSF4 isolate used to inoculate greenhouse *N. benthamiana* was originally isolated from *B. vulgaris* at the Taipei Botanical Garden in 1994 [17].

Selection of satBaMV P20 function

Among our satBaMV populations, the proportion of single nucleotide polymorphism (nucleotide substitutions that appear only once in the sample population) was relatively high, from 58% (Bo) to 79% (DIM) or to 88% when including the greenhouse (B+F4) population. The finding of the presence of excess rare variants is further corroborated with a more rigorous analysis with the Tajima’s *D* test [49], a statistical test designed to distinguish between neutral evolution of nucleotide sequences from non-neutral mechanisms, such as natural selection. All populations had

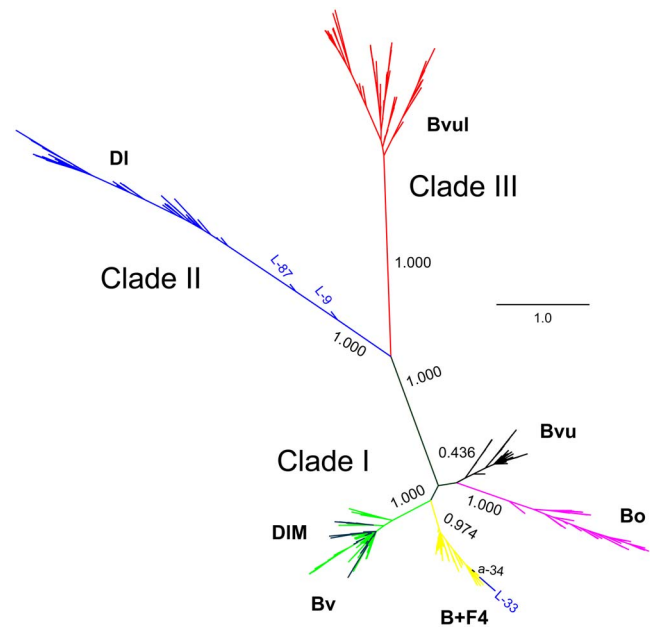


Figure 3. Phylogenetic relationship among different satBaMV populations. Unrooted Bayesian tree was constructed with full-length satBaMV genomic sequences. For clarity of presentation, sequence names are not shown, except for the listed (see text for details). Yellow, B+F4; magenta, Bo; green, Bv; black, Bvu; red, BvuI; blue, DI; purple, DIM. Posterior probabilities are shown for major branches. doi:10.1371/journal.pone.0108015.g003

significant negative *D* values, indicating the presence of excess rare variants (Table 2). Similar tests, such as the Fu and Li’s *D** and *F** [50] tests, showed the same pattern (data not shown).

Typically, population expansion and purifying selection are the two mechanisms commonly invoked to explain the finding of a negative Tajima’s *D* value [49], although these two explanations are not mutually exclusive. In our case, the sole protein-encoding gene *P20* in the satBaMV genome provided us the opportunity to test whether natural selection for *P20* function plays a role in the evolution of the satBaMV sequence. One of the commonly used methods for detecting the signatures of natural selection is the *dN/dS* test [55]. If the ratio (ω) is <1 , then it is customarily seen as evidence of purifying selection. However, if $\omega >1$, then it is indicative of positive selection favoring amino acid substitution. If $\omega =1$, then neither purifying selection nor positive selection is responsible for the observed nucleotide sequence pattern. With the commonly used *dN/dS* test implemented in MEGA [48], we found that 4 of the 7 populations (*i.e.*, Bvu, DI, DIM, and BvuI) showed significantly smaller *dN* values than corresponding *dS* values (Table 2), indicating the presence of purifying selection. Not all populations had significant differences between *dS* and *dN* values (*e.g.*, B+F4, Bo, and Bv), but all had a higher *dS* than *dN* value (Table 2). Together with the finding of negative Tajima’s *D* values, these results suggest a consistent explanation that purifying selection against less-fit *P20* variants are at least in part responsible for the evolution of satBaMV.

Discussion

Genome-size polymorphism and RT-PCR artifact

RT-PCR cloning followed by DNA sequencing is commonly used for population genetic studies of RNA viruses. However, the very act of RT-PCR can introduce sequence artifacts in the form

Table 2. Population genetic statistics and tests of satBaMV populations.

Population*	Bo	Bv	Bvu	DI	DIM	Bvul	B+F4
Tajima's D (p) [†]	-1.508 (0.042)	-2.629 (<0.001)	-2.230 (<0.001)	-1.530 (0.037)	-1.981 (0.007)	-1.466 (0.041)	-2.268 (<0.001)
Z (p) [‡]	1.278 (0.102)	1.189 (0.118)	2.094 (0.019)	2.315 (0.012)	2.094 (0.019)	2.474 (0.007)	0.284 (0.389)

*For population designation, see Figure 1 legend.

[†]Tajima's D test and associated probability in bracket. For Tajima's D test, $H_A = D_{obs} > D_{exp}$, with 100,000 replicates.

[‡] Z test and associated probability in bracket. $Z = dS-dN$ and the $H_A = dN < dS$, with 500 replicates.
doi:10.1371/journal.pone.0108015.t002

of point mutations, deletion/insertion [56,57], and sometimes recombination [58]. To minimize laboratory artifacts, we used a high-fidelity DNA polymerase for our amplifications. Nevertheless, we could not prevent the occurrence of deletion mutations in our samples, specifically the 732-nt variant, and possibly others as well. The deletion is possibly due to a secondary structure in the satBaMV genome encountered by the reverse transcriptase during the first-strand cDNA synthesis [57]. While the 732-nt variant may be an RT-PCR artifact, the phylogenetic relationship within this group still mirrored that of its full-length counterpart. For example, individual sequences from the same population still clustered together, with the same exception for those isolated from *D. latiflorus* Munro cv. "Mei-nung" in southern Taiwan and those from *B. ventricosa* McClure in Hainan Island, China, which formed a well-supported clade (data not shown). Furthermore, within each satBaMV population, the 732-nt sequences were interspersed among the full-length ones, which suggests rather than originating from a single or a few founding mutational events, most of the 732-nt sequences represent independent deletion events (data not shown). One possible explanation is that within a given satBaMV population, a portion of the genomes assumes a stable alternative secondary structure that, when encountered by a reverse transcriptase during RT-PCR, would facilitate the creation of a deletion mutation.

The 732-nt variant was demonstrated to be a laboratory artifact, but the remaining shorter genomes, especially those in both of the *D. latiflorus* Munro hosts (Table 1), may have a more biologically relevant origin. An *in planta* study with a satRNA associated with CMV (satCMV) [59] showed that the deletion mutation rate of the helper viral replicase depended on the plant host. The deletion rate was approximately 7-fold higher in pepper than tobacco. A similar dependency of the BaMV replicase on bamboo species could explain an increased proportion of the satBaMV genomes deleted when in *D. latiflorus* Munro.

Phylogenetic relationships among satBaMV populations

Our previous surveys of satBaMV in Taiwan revealed two readily discernible clades: one with only satBaMVs isolated from *D. latiflorus* Munro (but see below) and the other including all other samples, as well as some from *D. latiflorus* Munro [23,29]. We revealed a third exclusive clade composed of only individuals from *B. vulgaris* in India (Bvul) (Figure 3). When previous and current full-length satBaMV sequences were used in phylogenetic reconstruction, the same three main clades were retrieved (Figure 4). Because collection of these samples spanned more than a decade, the existence of these distinct clades suggests that they are quite stable in the satBaMV populations.

Clade II was composed of almost entirely individuals isolated from *D. latiflorus* Munro from various locations in Taiwan. Only two other bamboo species, *B. dolichoclada* cv. Stripe [29] and *G. levis* [23], were found to carry satBaMVs in this clade. However, at least eight different bamboo species were found infected with Clade I in Taiwan (Figure 5). Whether the paucity of host species for Clade II is due to biased sampling or genuine restricted host range is not known. Nevertheless, Clade II seems to have a narrower host range but approximately 7-fold higher genetic diversity (Figure 2), than Clade I. Clade I may have initially derived from the high-diversity Clade II, with subsequent host range expansion. Alternatively, Clade II may have derived from Clade I, accompanied by an increased mutation rate of the helper BaMV replicase in a different host, as occurred with the satCMV [59], thus maintaining a high steady-state nucleotide diversity in the population. So far, we cannot differentiate between these 2 scenarios. As for Clade III, which is only found in India with *B.*

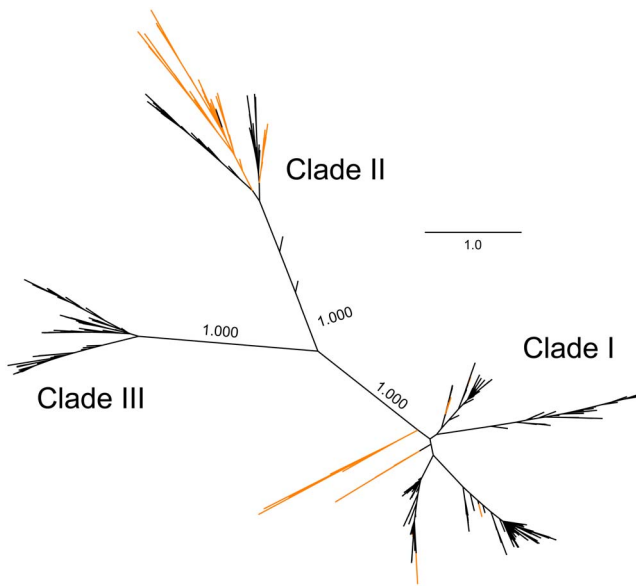


Figure 4. Phylogenetic relationship among all known full-length satBaMV. Unrooted Bayesian tree was constructed using full-length satBaMV genomic sequences from the current (black) and previous (orange) studies. Previous sequences used are from GenBank accession numbers AY205159–AY205231, except AY205176, AY205199, and AY205209. Posterior probabilities are shown for major branches. doi:10.1371/journal.pone.0108015.g004

vulgaris, our current sampling is too limited to determine whether this clade would have a narrow host range or not.

The existence of distinct satBaMV clades prompted us to examine the phylogenetic relationship of the helper BaMV isolates. Our preliminary analysis showed no immediately discernible clustering among the BaMV isolates (Figure 6), as compared with what is found for satBaMV. This preliminary result suggests that satRNAs, while obligatorily requiring the helper viral functions for replication, encapsidation, and efficient movement, could have very different evolutionary histories from their cognate helper viruses.

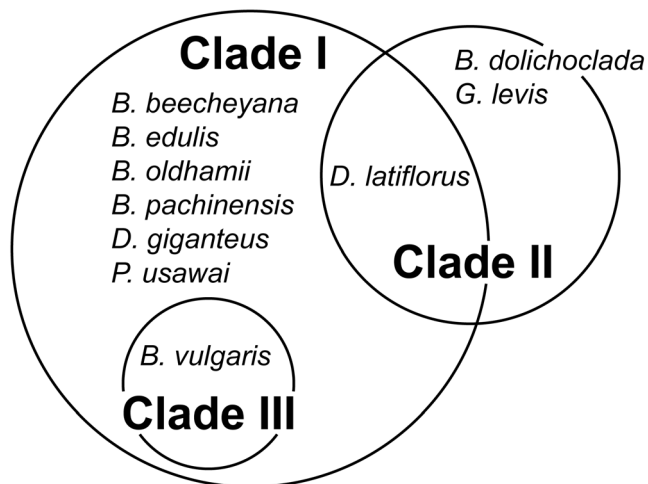


Figure 5. Distribution of bamboo species infected with different satBaMV clades. Information from both current and previous studies [23,29] were used to group bamboo species according to the satBaMV clade(s) found infecting them. doi:10.1371/journal.pone.0108015.g005

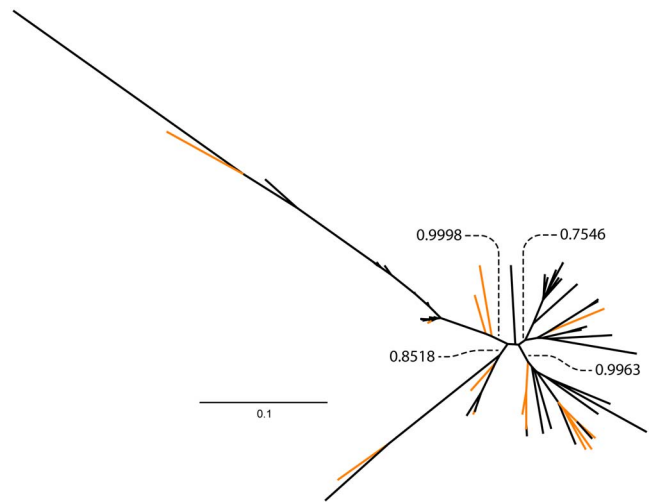


Figure 6. Phylogenetic relationship among BaMV isolates. Unrooted Bayesian tree was constructed using sequences encoding the RNA-dependent RNA polymerase from BaMV isolated from various bamboo species. Highlighted isolates (orange) denote BaMVs from the Ma bamboo (*D. latiflorus* Munro). Posterior probabilities (dashed lines) are shown for major branches. doi:10.1371/journal.pone.0108015.g006

Evolution of interference phenotype

Two prototypical isolates, BSL6 [29] and BSF4 [17], are commonly used in studies of satBaMV. The former interferes greatly with the replication of the helper BaMV [30], thus reducing the severity of disease symptoms, whereas the latter interferes only slightly, thus showing a typical mosaic symptom on co-infection with BaMV [19,30]. Competition for access of the BaMV-encoded replicase is thought to be the basis for replication interference [28,60]. One important determinant for interference is located at the AHSL at the 5' UTR, of which both the helper BaMV and satBaMV share a similar secondary structure [23,24]. Within the 5' AHSL, two nucleotide positions, 60 and 82 (based on the BSF4 sequence [17]), are critical for determining the interference phenotype [28]. If either position contains a U instead of C, the resulting satBaMV would not be able to (or only slightly) interfere with BaMV replication [28]. In our satBaMV populations, all sequences have a C at position 60 (or homologous positions), and almost all have a C at position 82 as well. The only exceptions are the satBaMV population from the greenhouse-infected *N. benthamiana*, for which 38.3% have a U in position 82, and one sequence (isolate a-34) from the *B. vulgaris* and three sequences (isolates L-9, L-33, and L-87) from *D. latiflorus* Munro that also have a U at the homologous position 82. The original sequence used for infecting *N. benthamiana* is BSF4, which contains a U at position 82 [17]. Therefore, even within a short time (*i.e.*, 7 dpi), we observed the evolution of U to C at position 82 from the initial 100% to 38.3%. Presumably, selection for accessing the replicase, against both the ancestral BSF4 and the helper BaMV, drives the evolution at this position. Interestingly, both sequences a-34 and L-33 are found clustered with the satBaMV population from the greenhouse-infected *N. benthamiana* (Figure 2), suggesting that satBaMV sequences very similar to BSF4, first reported in 1994 [17], still exist in the field and coexist with other satBaMV clades. Also, all sequences isolated from *D. latiflorus* Munro (DI) (except the L-33) clustered together, forming a well-supported Clade II, but both L-9 and L-87 seem to be distantly related to all the other sequences in this clade (Figure 3).

These two sequences may represent an intermediate stage between Clade I and II satBaMVs.

With our current phylogeny, BSF4 belongs to Clade I and BSL6 to Clade II. However, mapping of the nucleotide state at positions 60 and 82 showed that, in the field, both clades are dominated by sequences that may greatly interfere with BaMV replication. The same pattern was also found with individuals in Clade III. Most of the satBaMV sequences found in the field probably also have a C in positions that are homologous to position 82. Other features of the genome, such as the AHSL structure and the surrounding sequence in the 5' UTR region, could influence the interference phenotype, but the evolution of these satBaMV clades is unlikely to be driven by the evolution of the interference phenotype.

Selection of the P20 protein

The satBaMV-encoded P20 protein is important in long-distance movement through the phloem [17,21,22]. P20 plays a supporting role in satBaMV replication as well. For example, none of the truncated P20 mutants could reach the same accumulation level as that of the wild-type P20 in the protoplast assay [19]. Furthermore, other satRNA-encoded proteins are required for satRNA replication [61–63]. Therefore, P20 likely facilitates its replication as well as long-distance movement within the infected host.

Our analysis suggests that the *P20* gene is under purifying selection against less-fit P20 proteins, which, presumably, are not efficient in facilitating satBaMV movement within the host plant and/or its own replication. However, for purifying selection to be the main driving force for the evolution of satBaMV, the interaction between individual satBaMV genomes and the P20 proteins inside an infected cell needs to be predominantly *cis* in nature; that is, the genome that sustained a deleterious mutation in the *P20* gene should, on average, “pay” a fitness cost for expressing a less fit P20 protein, either as a result of less-efficient long-distance and/or less-efficient replication of the satBaMV genome. However, if P20 is a diffusible protein that interacts with any satBaMV genomes for movement and/or replication (*i.e.*, a mainly *trans*-acting interaction), then other fitter P20 proteins can still facilitate the movement and/or replication of the satBaMV genome that carries a less-fit *P20* mutation. The within-cell complementation of fit P20 proteins with mutant satBaMV genomes could greatly reduce or completely abolish the effect of purifying selection.

The replication of the potexviruses, of which BaMV is a member, may provide clues to whether the prerequisite *cis* interaction is possible. Replication of many positive-strand plant

RNA viruses is accompanied by extensive intracellular membrane reorganization, usually involving the formation of viral-induced membrane vesicles or invagination [64]. One consequence of RNA genome compartmentalization is the almost exclusive interaction between the protein products and the genomes expressing them, effectively a *cis* interaction. Whether BaMV replication, and by extension satBaMV replication, also involves rearranged membrane structures is unclear. Although other potexviruses [65] adopt such a strategy for genome replication. In the case of satBaMV, it would be interesting to know if co-infection of a wild-type satBaMV with a P20 mutant, such as the one with the P20 gene replaced by *cat* [19], may result in the translocation of the mutant satBaMV, thus establishing the *cis*-acting nature between P20 and the genome expressing it.

Supporting Information

Figure S1 Genomic deletions in satBaMV isolates. The prototypical satBaMV genome of the BSF4 isolate (GenBank accession no. AY205227) was used to illustrate region(s) of deletion in the genomes of the most common isolate in each “short” category listed in Table 1. Open bar shows the encoded *P20* gene and hatched bars the regions of deleted genome. Numbers with vertical bars indicate the nucleotide positions (of BSF4) for the start and end of the *P20* gene or the deleted regions. Numbers at the end of each genome show the specific genome length. Various single-nucleotide deletions, found in some genomes, are not shown in the illustration.

(DOCX)

Acknowledgments

We thank Drs. Andrew O. Jackson (University of California, Berkeley, CA, USA) and Jia-Lin Yu (China Agricultural University, Beijing, China) for their help in collecting BaMV-infected bamboo from Hainan Island, China and Adam Bell for preparing high-resolution figures for publication. This study was supported by grants from the Institute of Plant and Microbial Biology (IPMB), Academia Sinica, Taiwan, to NSL. INW received funding from the National Science Council, Taiwan (NSC 101-2811-B-001-024), IPMB, and Academia Sinica during his sabbatical visit.

Author Contributions

Conceived and designed the experiments: YHH NSL. Performed the experiments: CWL SMY. Analyzed the data: INW CCH WBY. Contributed reagents/materials/analysis tools: CWL WBY YHH NSL. Wrote the paper: INW NSL.

References

- Hu CC, Hsu YH, Lin NS (2009) Satellite RNAs and Satellite Viruses of Plants. *Viruses* 1: 1325–1350.
- Roossinck MJ, Sleat D, Palukaitis P (1992) Satellite RNAs of plant viruses: structures and biological effects. *Microbiol Rev* 56: 265–279.
- Simon AE, Roossinck MJ, Havelda Z (2004) Plant virus satellite and defective interfering RNAs: new paradigms for a new century. *Annu Rev Phytopathol* 42: 415–437.
- Murant AF, Mayo MA (1982) Satellites of plant viruses. *Annual Review of Phytopathology* 20: 49–70.
- Tien P, Wu GS (1991) Satellite RNA for the biocontrol of plant disease. *Adv Virus Res* 39: 321–339.
- Nee S (2000) Mutualism, parasitism and competition in the evolution of coviruses. *Philos Trans R Soc Lond B Biol Sci* 355: 1607–1613.
- Harrison BD, Mayo MA, Baulcombe DC (1987) Virus resistance in transgenic plants that express cucumber mosaic virus satellite RNA. *Nature* 328: 799–802.
- Gerlach WL, Llewellyn D, Haseloff J (1987) Construction of a plant disease resistance gene from the satellite RNA of tobacco ringspot virus. *Nature* 328: 802–805.
- Lin KY, Hsu YH, Chen HC, Lin NS (2013) Transgenic resistance to Bamboo mosaic virus by expression of interfering satellite RNA. *Mol Plant Pathol* 14: 693–707.
- Kong Q, Wang J, Simon AE (1997) Satellite RNA-mediated resistance to Turnip crinkle virus in *Arabidopsis* involves a reduction in virus movement. *Plant Cell* 9: 2051–2063.
- Collmer CW, Howell SH (1992) Role of satellite RNA in the expression of symptoms caused by plant viruses. *Annu Rev Phytopathol* 30: 419–442.
- Sleat DE, Palukaitis P (1990) Site-directed mutagenesis of a plant viral satellite RNA changes its phenotype from ameliorative to necrogenic. *Proc Natl Acad Sci USA* 87: 2946–2950.
- Roossinck MJ (2001) Cucumber mosaic virus, a model for RNA virus evolution. *Mol Plant Pathol* 2: 59–63.
- Lin MT, Kitajima EW, Cupertino FP, Costa CL (1977) Partial purification and some properties of bamboo mosaic virus. *Phytopathology* 67: 1439–1443.
- Lin N-S, Lin F-Z, Huang T-Y, Hsu Y-H (1992) Genome properties of bamboo mosaic virus. *Phytopathology* 82: 731–734.
- Lin NS, Lin BY, Lo NW, Hu CC, Chow TY, et al. (1994) Nucleotide sequence of the genomic RNA of bamboo mosaic potexvirus. *J Gen Virol* 75: 2513–2518.

17. Lin NS, Hsu YH (1994) A satellite RNA associated with bamboo mosaic potexvirus. *Virology* 202: 707–714.
18. Liu JS, Lin NS (1995) Satellite RNA associated with bamboo mosaic potexvirus shares similarity with satellites associated with sobemoviruses. *Arch Virol* 140: 1511–1514.
19. Lin NS, Lee YS, Lin BY, Lee CW, Hsu YH (1996) The open reading frame of bamboo mosaic potexvirus satellite RNA is not essential for its replication and can be replaced with a bacterial gene. *Proc Natl Acad Sci USA* 93: 3138–3142.
20. Tsai MS, Hsu YH, Lin NS (1999) Bamboo mosaic potexvirus satellite RNA (satBaMV RNA)-encoded P20 protein preferentially binds to satBaMV RNA. *J Virol* 73: 3032–3039.
21. Palani PV, Kasiviswanathan V, Chen JC, Chen W, Hsu YH, et al. (2006) The arginine-rich motif of Bamboo mosaic virus satellite RNA-encoded P20 mediates self-interaction, intracellular targeting, and cell-to-cell movement. *Mol Plant-Microbe Interact* 19: 758–767.
22. Palani PV, Chen JC, Liou MR, Chen HC, Hsu YH, et al. (2012) Phosphorylation of bamboo mosaic virus satellite RNA (satBaMV)-encoded protein P20 downregulates the formation of satBaMV-P20 ribonucleoprotein complex. *Nuc Acids Res* 40: 638–649.
23. Yeh WB, Hsu YH, Chen HC, Lin NS (2004) A conserved secondary structure in the hypervariable region at the 5′ end of Bamboo mosaic virus satellite RNA is functionally interchangeable. *Virology* 330: 105–115.
24. Chen SC, Desprez A, Olsthoorn RC (2010) Structural homology between bamboo mosaic virus and its satellite RNAs in the 5′ untranslated region. *J Gen Virol* 91: 782–787.
25. Huang YW, Hu CC, Lin CA, Liu YP, Tsai CH, et al. (2009) Structural and functional analyses of the 3′ untranslated region of Bamboo mosaic virus satellite RNA. *Virology* 386: 139–153.
26. Hsu YH, Chen HC, Cheng J, Annamalai P, Lin BY, et al. (2006) Crucial role of the 5′ conserved structure of bamboo mosaic virus satellite RNA in downregulation of helper viral RNA replication. *J Virol* 80: 2566–2574.
27. Chen HC, Hsu YH, Lin NS (2007) Downregulation of Bamboo mosaic virus replication requires the 5′ apical hairpin stem loop structure and sequence of satellite RNA. *Virology* 365: 271–284.
28. Chen HC, Kong LR, Yeh TY, Cheng CP, Hsu YH, et al. (2012) The conserved 5′ apical hairpin stem loops of bamboo mosaic virus and its satellite RNA contribute to replication competence. *Nuc Acids Res* 40: 4641–4652.
29. Liu JS, Hsu YH, Huang TY, Lin NS (1997) Molecular evolution and phylogeny of satellite RNA associated with bamboo mosaic potexvirus. *J Mol Evol* 44: 207–213.
30. Hsu Y-H, Lee Y-S, Liu J-S, Lin N-S (1998) Differential interactions of bamboo mosaic potexvirus satellite RNAs, helper virus, and host plants. *Mol Plant-Microbe Interact* 11: 1207–1213.
31. Kurath G, Palukaitis P (1989) RNA sequence heterogeneity in natural populations of three satellite RNAs of cucumber mosaic virus. *Virology* 173: 231–240.
32. Aranda MA, Fraile A, Garcia-Arenal F (1993) Genetic variability and evolution of the satellite RNA of cucumber mosaic virus during natural epidemics. *J Virol* 67: 5896–5901.
33. Alonso-Prados JL, Aranda MA, Malpica JM, Garcia-Arenal F, Fraile A (1998) Satellite RNA of cucumber mosaic cucumovirus spreads epidemically in natural populations of its helper virus. *Phytopathology* 88: 520–524.
34. Garcia-Arenal F, Escriu F, Aranda MA, Alonso-Prados JL, Malpica JM, et al. (2000) Molecular epidemiology of Cucumber mosaic virus and its satellite RNA. *Virus Res* 71: 1–8.
35. Lin N-S, Chai Y-J, Huang T-Y (1993) Incidence of bamboo mosaic potexvirus in Taiwan. *Plant Disease* 77: 448–450.
36. Lin MK, Chang BY, Liao JT, Lin NS, Hsu YH (2004) Arg-16 and Arg-21 in the N-terminal region of the triple-gene-block protein 1 of Bamboo mosaic virus are essential for virus movement. *J Gen Virol* 85: 251–259.
37. Vervoerd TC, Dekker BM, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. *Nuc Acids Res* 17: 2362.
38. Shi BJ, Ding SW, Symons RH (1997) Plasmid vector for cloning infectious cDNAs from plant RNA viruses: high infectivity of cDNA clones of tomato aspermy cucumovirus. *J Gen Virol* 78: 1181–1185.
39. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7: 539.
40. Martin D, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16: 562–563.
41. Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology* 265: 218–225.
42. Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* 21: 98–102.
43. Maynard Smith J (1992) Analyzing the mosaic structure of genes. *J Mol Evol* 34: 126–129.
44. Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci USA* 98: 13757–13762.
45. Gibbs MJ, Armstrong JS, Gibbs AJ (2000) Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16: 573–582.
46. Boni MF, Posada D, Feldman MW (2007) An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 176: 1035–1047.
47. Martin DP, Lemey P, Lott M, Moulton V, Posada D, et al. (2010) RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26: 2462–2463.
48. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
49. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
50. Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
51. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Syst Biol* 61: 1–4.
52. Rambaut A (2009) FigTree v1.3.1: Tree Figure Drawing Tool.
53. Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76: 5269–5273.
54. Sokal RR, Rohlf FJ (2011) *Biometry*. New York: W. H. Freeman.
55. Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335: 167–170.
56. Bracho MA, Moya A, Barrio E (1998) Contribution of *Taq* polymerase-induced errors to the estimation of RNA virus diversity. *J Gen Virol* 79: 2921–2928.
57. Domingo-Calap P, Sentandreu V, Bracho MA, Gonzalez-Candelas F, Moya A, et al. (2009) Unequal distribution of RT-PCR artifacts along the E1–E2 region of *Hepatitis C virus*. *J Virol Methods* 161: 136–140.
58. Mader RM, Schmidt WM, Sedivy R, Rizovski B, Braun J, et al. (2001) Reverse transcriptase template switching during reverse transcriptase-polymerase chain reaction: artificial generation of deletions in ribonucleotide reductase mRNA. *J Lab Clin Med* 137: 422–428.
59. Pita JS, de Miranda JR, Schneider WL, Roossinck MJ (2007) Environment determines fidelity for an RNA virus replicase. *J Virol* 81: 9072–9077.
60. Huang YW, Hu CC, Liou MR, Chang BY, Tsai CH, et al. (2012) Hsp90 interacts specifically with viral RNA and differentially regulates replication initiation of Bamboo mosaic virus and associated satellite RNA. *PLoS Pathog* 8: e1002726.
61. Liu YY, Cooper JI (1993) The multiplication in plants of arabis mosaic virus satellite RNA requires the encoded protein. *J Gen Virol* 74: 1471–1474.
62. Hemmer O, Oncino C, Fritsch C (1993) Efficient replication of the in vitro transcripts from cloned cDNA of tomato black ring virus satellite RNA requires the 48 K satellite RNA-encoded protein. *Virology* 194: 800–806.
63. Hans F, Pinck M, Pinck L (1993) Location of the replication determinants of the satellite RNA associated with grapevine fanleaf nepovirus (strain F13). *Biochimie* 75: 597–603.
64. den Boon JA, Ahlquist P (2010) Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu Rev Microbiol* 64: 241–256.
65. Verchot-Lubicz J, Ye CM, Bamunusinghe D (2007) Molecular biology of potexviruses: recent advances. *J Gen Virol* 88: 1643–1655.