RNA silencing suppression activity encoded by plant-infecting DNA satellites from Africa

Mi Zhou
RNA silencing suppression activity encoded by plant-infecting DNA satellites from Africa

Mi Zhou

Supervisor: Prof. Anders Kvarnheden
Assistant Supervisor: Muhammad Naeem Sattar
Examiner: Prof. Johan Meijer

Credits: 30 hec
Level: Advanced D
Course title: Independent project in Biology
Course code: EX0564
Biotechnology-Master’s programme

Place of publication: Uppsala
Year of publication: 2012
Serienamn: Examensarbete / Institutionen för växtbiologi och skogsgenetik, SLU
ISSN: 1651-5196 Nr:131
Online publication: http://stud.epsilon.slu.se

Key Words: alphasatellite, betasatellite, begomovirus, Alpha-Rep, βC1, RNA interference, suppressor activity, phylogenetic analyses

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences
Faculty of Natural Resources and Agricultural sciences
Department of Plant Biology and Forest Genetics
Begomoviruses are whitefly-transmitted viruses, which can induce severe diseases in various vegetable, fiber and cereal crops around the world. Begomoviruses have a circular single-stranded (ss) DNA genome, and they often infect plants together with other circular ssDNA molecules, which are the so-called alphasatellites and betasatellites. Alpha-Rep is a replication initiator protein encoded by the genome of alphasatellites. RNA silencing is an important part of plants’ natural defense against viruses. Plant viruses and their associated satellites adopt different strategies to overcome this defense mechanism and promote viral infection. The aim of this project was to test whether the Alpha-Rep proteins of two alphasatellites (Okra yellow crinkle Cameroon alphasatellite, OYCrCMA; Ageratum leaf curl Cameroon alphasatellite, ALCCMA) can suppress RNA silencing. Gene cloning technology was used to express these two Alpha-Rep proteins in RNA silencing-induced plants. The βC1 protein specified by betasatellites has been reported to have strong suppression activity of RNA silencing. The suppressor activity of βC1 encoded by Cotton leaf curl Gezira betasatellite (CLCuGB) was investigated mostly as a positive control. The Alpha-Rep genes of OYCrCMA and ALCCMA and the βC1 gene of CLCuGB were first cloned into the intermediate pJET1.2 vector, and then transferred to the binary vector PLH7000. To test the ability of the satellite proteins to suppress RNA silencing, a PA-GFP construct (to express the GFP transgene) and an RNAi-GFP construct (to induce RNA silencing) were infiltrated into whole leaves of wildtype Nicotiana benthamiana plants using Agrobacterium tumefaciens at the 4 leaf stage. At 4 dpi (days post inoculation), GFP-silenced leaves were inoculated with Agrobacterium containing the cloned satellite genes. Visual detection of green fluorescence in the N. benthamiana plants infiltrated with distinct constructs was performed using ultraviolet (UV) illumination. The results showed that both of the Alpha-Rep proteins can suppress RNA silencing, and that the Alpha-Rep protein of ALCCMA has much stronger suppression activity than that of OYCrCMA.
Popular science

Alphasatellites are circular DNA molecules infecting plants together with some plant viruses. They express proteins called Alpha-Rep. Plants can adopt different strategies to overcome viral attack, one of which is RNA silencing. Plant viruses are both targets and inducers of this mechanism. The genetic form of dsRNA (double stranded RNA) induces RNA silencing which specifically degrades viral RNA through the mediation of siRNA (short interfering RNA). This project was to investigate whether Alpha-Rep proteins from two alphasatellites have the ability to support viral attack by suppressing RNA silencing. The two alphasatellites are Okra yellow crinkle Cameroon alphasatellite and Ageratum leaf curl Cameroon alphasatellite, and both of them are from Africa. Through gene cloning technology, these two proteins were expressed in RNA-silencing activated plants. By direct observation of the phenomenon of RNA silencing, the result showed that both of the Alpha-Rep proteins can suppress RNA silencing, and that the Alpha-Rep protein of Ageratum leaf curl Cameroon alphasatellite has much stronger activity than that of Okra yellow crinkle Cameroon alphasatellite.
Contents

Introduction ..................................................................................................................5
Family *Geminiviridae* ....................................................................................................5
Genus *Begomovirus* ......................................................................................................6
Alphasatellites ................................................................................................................7
Betasatellites ..................................................................................................................8
RNA silencing ................................................................................................................8
Gene cloning ....................................................................................................................10
Aim of this study .............................................................................................................11
Materials and Methods .................................................................................................12
  Gene amplification by polymerase chain reaction (PCR) ............................................12
  Production of gene expression constructs for RNA silencing suppression analysis .......13
  Sequence analysis and construction of phylogenetic trees .........................................14
  Plant inoculation .........................................................................................................14
  Imaging and collection of samples .............................................................................15
Results ..........................................................................................................................16
  Gene cloning of OYCrCMA-Rep, ALCCMA-Rep and CLCuGB-βC1 in pJET1.2 ..........16
  Sequencing and bioinformatics analysis .....................................................................19
  Further cloning of the Alpha-Rep genes in PLH7000 .................................................23
  RNA silencing suppression analysis ..........................................................................25
Discussion ......................................................................................................................27
References ......................................................................................................................29
Acknowledgements .......................................................................................................33
Introduction

Family Geminiviridae

Viruses of the family Geminiviridae have DNA genomes and replicate via double-stranded DNA (dsDNA) intermediates. Geminiviruses are a diverse and large group of angiosperm-infecting viruses characterized by a twinned icosahedral capsid and circular single-stranded DNA (ssDNA) genomes contained in a capsid\(^1,2\). Geminiviruses can induce severe diseases in various vegetable, fiber and cereal crops around the world, including cassava, squash, cotton, bean, tobacco, pepper, tomato, sugarcane, wheat and maize. In addition to their role as plant pathogens, geminiviruses are perfect model systems for studying plant genome transcription and replication due to their small genome size, simple genome structure and their reliance on the molecular machinery of host cells\(^2,3\).

The viruses belonging to the family Geminiviridae are classified into four genera (Begomovirus, Curtovirus, Topocuvirus, and Mastrevirus) based on their genome organization, host range and insect vector (Table 1)\(^4\).

<table>
<thead>
<tr>
<th>Genus</th>
<th>Type species</th>
<th>Genomic organization</th>
<th>Plant range</th>
<th>Insect vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastrevirus</td>
<td>Maize streak virus</td>
<td>Monopartite</td>
<td>Monocots/dicots</td>
<td>Leafhoppers</td>
</tr>
<tr>
<td>Topocuvirus</td>
<td>Tomato pseudo curly top virus</td>
<td>Monopartite</td>
<td>Dicots</td>
<td>Treehoppers</td>
</tr>
<tr>
<td>Curtovirus</td>
<td>Beet curly top virus</td>
<td>Monopartite</td>
<td>Dicots</td>
<td>Leafhoppers, Treehoppers</td>
</tr>
<tr>
<td>Begomovirus</td>
<td>Bean golden mosaic virus</td>
<td>Mono- or bipartite</td>
<td>Dicots</td>
<td>Whiteflies</td>
</tr>
</tbody>
</table>

Table 1. Classification of the family Geminiviridae.
Genus *Begomovirus*

The genus *Begomovirus*, which is the most widely distributed and diverse genus in the family *Geminiviridae*, has different geographic lineages. There are two major clusters: the Old World cluster (OW) for the viruses from Asia, Australia, Africa and Europe and the New World cluster (NW) for the viruses from America\(^5\). Begomoviruses are transmitted by the whitefly *Bemisia tabaci*. Typical begomovirus species include *Bean golden mosaic virus* (BGMV), *African cassava mosaic virus* (ACMV), *Pepper huasteco virus* (PHV), *Squash leaf curl virus* (SqLCV), *Tomato golden mosaic virus* (TGMV), *Tomato leaf curl virus* (TLCV), and *Tomato yellow leaf curl virus* (TYLCV). Begomoviruses are either monopartite (a single genome component) or bipartite (two genome components), and they often infect plants together with other circular ssDNA molecules, which are the so-called alphasatellites and betasatellites\(^5\).

The genome of monopartite begomoviruses contains 6 overlapping genes, which encode all the necessary proteins needed in the process of encapsidation, replication, transcription and virus movement within or between plants. All the 6 genes are arranged in two clusters separated by an intergenic region (IR). One cluster of genes arranged in the clockwise direction is designated as virion sense because the virion sense strand has the same sequence as the corresponding mRNAs. The template DNA strand is complementary to the ssDNA encapsidated in the viral capsid and it exists only in the dsDNA intermediate. Similarly, the other cluster of genes in the counter-clockwise direction is designated as complementary sense because the DNA strand used as a transcriptional template is just the same as the one encapsidated in the viral capsid. Generally, the proteins with function in viral transcription and replication such as C1 (replication initiator protein), C2 (transcription activator protein) and C3 (replication enhancer protein) are encoded in the complementary sense. Similarly, proteins involved in virus movement and encapsidation, V1 (coat protein), V2 (pre-coat protein) and C4 (an uncharacterized protein) are encoded in the virion sense\(^6,7\).

The genome of bipartite begomoviruses is composed of two circular ssDNA molecules called the DNA-A component and DNA-B component, respectively, both of which are necessary for the infection process. The sequences also have two clusters of genes separated by an intergenic region (IR). This IR contains a region highly conserved between DNA-A and DNA-B called the common region (CR). All the cis-elements necessary for viral replication exist in this CR. For a typical bipartite begomovirus, DNA-A contains four to five genes. There is only one gene (CP) arranged in virion sense, while the genes for Rep, TrAP, and REn are arranged in the complementary-sense direction. DNA-B contains only two genes (BV1 in the virion sense and BC1 in the
complementary sense), both of which encode proteins responsible for movement of the viral DNA and that are necessary for symptom production and systemic movement\(^8,9\).

**Alphasatellites**

The begomovirus-associated alphasatellites are satellite-like self-replicating molecules which depend on the helper virus for whitefly transmission, encapsidation and movement. Almost all the identified begomovirus-associated alphasatellites are associated with OW monopartite begomovirus diseases such as ageratum yellow vein disease in South East Asia, tomato leaf curl disease from China and cotton leaf curl disease from India and Pakistan. However, recently an atypical alphasatellite (Melon chlorotic mosaic alphasatellite; MeCM) has been found to be associated with a bipartite begomovirus (*Melon chlorotic mosaic virus*; MeCMV) in the NW\(^10\). Alphasatellites are not essential for begomovirus-betasatellite infection, and infection of alphasatellites does not affect symptom development. The genome structure of alphasatellites is highly conserved. It contains only one large gene which encodes the alpha-replication initiator protein (Alpha-Rep) of 36 kDa, an adenine-rich (A-rich) region and a hairpin structure (**Fig. 1**\(^11\)). Alphasatellites can maintain the suppression activity of post-transcriptional gene silencing (PTGS) in the newly developed tissue, but this suppression activity is only transient. At least two Alpha-Rep proteins encoded by *Gossypium darwinii* symptomless alphasatellite (GDarSLA) and *Gossypium mustelinium* symptomless alphasatellite (GMusSLA) have been shown to have suppression activity of RNA silencing. This capacity of Alpha-Rep can possibly explain the selective advantage of alphasatellite-associated begomoviruses in the process of evolution\(^12,13\). However, the knowledge about the silencing activity of Alpha-Rep is still very limited. In this study, we investigated two Rep proteins encoded by Okra yellow crinkle Cameroon alphasatellite (OYCrCMA) and Ageratum leaf curl Cameroon alphasatellite (ALCCMA) identified during the characterization of begomoviruses and satellites from Cameroon that infect the weed *Ageratum conyzoides* and the crop plant okra. OYCrCMA is a typical new alphasatellite, and ALCCMA is a typical old alphasatellite. They are evolutionarily distant away from each other, but both of them are from the OW. Investigation of RNA silencing suppression activity of OYCrCMA and ALCCMA could improve our knowledge about the two different types of alphasatellites from the OW\(^14\).
Figure 1. The genome organization of a typical alphasatellite.

Betasatellites

The begomovirus-associated betasatellites are pathogenicity determinants and associated to distinct plant diseases caused by monopartite begomoviruses from the OW. They entirely depend on the helper virus for their movement, transmission, encapsidation and replication. Betasatellites are specifically associated with their helper viruses regardless of geographical distribution and hosts. The genome of betasatellites contains only one gene encoding the βC1 protein of 13 kDa in the complementary sense. Also, a SCR (satellite conserved region) of around 220 nucleotides and an adenine-rich region of around 240 nucleotides are found to be present in the genome of betasatellites. The SCR is a genomic region highly conserved in all betasatellites15-18. The βC1 protein specified by betasatellites has been reported to have strong suppression activity of RNA silencing19. In this study, the suppressor activity of βC1 encoded by Cotton leaf curl Gezira betasatellite (CLCuGB) was investigated mostly as a positive control20.

RNA silencing

RNA silencing is an important part of plants’ natural defense against viruses21. The basic mechanism of RNA silencing is as follows. The target of RNA silencing is dsRNA, so mRNA molecules in eukaryotic cells are not recognized by this mechanism under normal conditions. Viruses are targets for RNA silencing because their genomes are either dsRNA originally or form dsRNA molecules during their infection cycles. Once dsRNA molecules are
formed, they are identified by specific binding proteins, which provide attachment positions for Dicer-like proteins. Dicers can cleave the dsRNA to siRNA (short interfering RNA), which is 20-30 nucleotides long. Subsequently, these siRNA molecules can be recognized by RISC (RNA-induced silencing complex) and separated into two sets of ssRNA. Only one set of ssRNA molecules complementary to the viral mRNA can be employed as a probe and attach to target mRNA to form a small region of dsRNA. These regions are recognition sites for RISC, which induces the degradation of viral mRNA (Fig. 2)\textsuperscript{22-25}. Although begomoviruses have DNA genomes and do not replicate using dsRNA intermediates, they have been proven to be targets of RNA interference. One possible reason is that dsRNA is formed in the process of bidirectional transcription. Five proteins encoded by begomoviruses and their satellites, including V2, C2, C4, BetaC1 and Alpha-Rep, possess RNA silencing suppressor activity\textsuperscript{26, 27}. Compared with the other four proteins, the alpha-Rep protein has shown the strongest suppressor activity\textsuperscript{12}.

\textbf{Figure 2.} General mechanism of RNA silencing. Main steps of RNA silencing: 1) Recognition of double-stranded RNA by Dicer, 2) Cleavage into short interfering RNA, 3) Formation of RNA induced silencing complex (RISC), 4) Degradation of target mRNA.
Gene cloning

Gene cloning refers to a methodology of DNA manipulation using endonucleases and ligases, from which a huge amount of DNA containing a single target gene can be obtained\(^2\). The general procedure for gene cloning comprises several steps. A gene of interest with restriction sites flanking both the 5' and 3' ends is cut by the corresponding endonucleases. Plasmids, circular DNA that can replicate autonomously within bacteria, are treated with the same endonucleases so that both the DNA fragments that contain the target gene and the plasmids have the same sticky ends complementary to each other. After mixing the DNA fragments and the vectors, ligases are used to amalgamate the target gene with the plasmid. Since several different recombinant products will be generated in the process of ligation, plasmids carrying the target gene must be selected and identified from the mixture. This can be achieved through several methods such as Lac selection and colony PCR. The recombinant plasmids are transformed into competent bacteria, in which they replicate within the bacterial cells and become distributed to the daughter cells during cell division. In most cases, only a small proportion of bacterial cells are capable of taking up a plasmid, so selection of transformants is an indispensable step. An antibiotic-resistance gene in the vector is often employed to achieve the selection. A colony with recombinant plasmids is obtained, in which every bacterium contains many copies of the target gene. Ultimately, the recombinant plasmids are extracted from the colony or a bacterial culture derived from a single colony\(^2\).

The invention of gene cloning promoted the great development of molecular biology, because this technology made it possible to isolate DNA containing a target gene in microgram quantities, which is sufficient for subsequent gene analyses\(^2\).
Aim of this study

To test whether two distantly related alphasatellites (Okra yellow crinkle Cameroon alphasatellite and Ageratum leaf curl Cameroon alphasatellite) have a Rep protein with RNA silencing suppression activity.
Materials and Methods

Gene amplification by polymerase chain reaction (PCR)

Satellite clones (OYCrcMA, GenBank accession number FN675284; ALCCMA, GenBank accession number FR717142; CLCuGB, GenBank accession number FM164725) in pBluescript-II KS+ were used as templates for PCR amplification. The coding sequences of the Alpha-Rep genes in OYCrcMA and ALCCMA, and the βC1 gene in CLCuGB were amplified by PCR with specific primer pairs. Restriction enzyme recognition sites, Ncol and Xbal, were included in the primers (Table 2). Together with the Ncol recognition site, a start codon was introduced in each forward primer, so the cloned genes will be transcribed from the new start codon instead of the original ones and the nucleotide following the new start codon in the Ncol recognition site will lead to frameshift mutation. In order to avoid any frameshift mutation, two additional nucleotides were added to form a new codon (GCC) following the start codon. Negative controls for PCR were run with water instead of the templates. The total reaction volume for PCR was 25 µL including 14.7 µL water, 5 µL Phusion enzyme reaction buffer (5×), 1 µL dNTP, 1.5 µL forward primer (10 µM), 1.5 µL reverse primer (10 µM), 0.3 µL Phusion enzyme (2 units/µL) and 1 µL template (50 ng). The PCR program used in this study started with heating the samples to 98°C for 1 minute followed by 34 PCR cycles. Each cycle included 30 seconds at 98°C for denaturing the template, 1 minute at 52°C for primer annealing and 1 minute at 72°C for elongation. Final elongation was carried out at 72°C for 10 minutes. Subsequently, the PCR samples were cooled down to 12°C to prevent further reactions.
Table 2. Primers used in the PCR amplification. Underlined sequences are restriction sites introduced to the 5’ end of the six primers. “F” in the names of the primers refers to forward primers and “R” represents reverse primers. The three nucleotides in red are the new codon following the start codon introduced to avoid frame shift mutation.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OYCrCMA-RepF</td>
<td>GTCATGCCCTTCTTATTAAGAGTGGTTTG</td>
</tr>
<tr>
<td>OYCrCMA-RepR</td>
<td>GTCCTTCTAAAGACAAATCGATTAATTT</td>
</tr>
<tr>
<td>ALCCMA-RepF</td>
<td>GTCATGCCCTTCTGTGCAATCTAGT</td>
</tr>
<tr>
<td>ALCCMA-RepR</td>
<td>GTCCTTCTAAATACCTACATCATC</td>
</tr>
<tr>
<td>CLCuGB-βC1F</td>
<td>GTCATGGCCACTATCAGTCAGGAGC</td>
</tr>
<tr>
<td>CLCuGB-βC1R</td>
<td>GTCCTAGTTACACGGATGATTATTATG</td>
</tr>
</tbody>
</table>

Production of gene expression constructs for RNA silencing suppression analysis

The PCR products were extracted from the gel using the GeneJET™ PCR purification Kit (Fermentas) and ligated into pJET1.2. The recombinant constructs were transformed into competent Escherichia coli cells (DH5α). Colony PCR was performed to check if the recombinant constructs were successfully transformed into the E. coli cells by amplifying the corresponding genes with the pJET1.2 sequencing primers. Recombinant plasmid was extracted from the E. coli cell cultures using a plasmid miniprep kit (Fermentas). Subsequently, the cloned Alpha-Rep genes were transferred from the vector pJET1.2 to the binary vector PLH7000 using the corresponding restriction endonucleases. The PA-GFP construct was used to express the GFP protein in wildtype Nicotiana benthamiana plants, and the construct RNAi-GFP was used to induce RNA silencing of GFP. The RNAi-GFP construct contains the first 200 bp of the GFP gene in the sense orientation, the same sequence in the antisense orientation, and an artificial intron separating the two GFP fragments (Fig. 3). When this DNA region is transcribed, it will form a hairpin structure containing dsRNA which will induce RNA silencing of GFP. The 35S-HC-Pro construct containing the HC-Pro gene from Potato virus Y (PVY) was employed as a positive control in this
study, since HC-Pro has been shown to be a very strong suppressor of RNA silencing\textsuperscript{36}. All the six constructs were transformed into \textit{Agrobacterium tumefaciens} GV3101 by electroporation\textsuperscript{36}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3}
\caption{The functional region of the RNAi-GFP construct}
\end{figure}

**Sequence analysis and construction of phylogenetic trees**

The three cloned genes in pJET1.2 were sent to Macrogen Inc., Korea for automated sequencing in forward direction. The nucleotide sequences were translated into the amino acid sequences using MEGA5\textsuperscript{37}. BLASTp was used to compare the amino acid sequence of the cloned genes with other sequences in GenBank\textsuperscript{38}. To search for mutations and construct phylogenetic trees, Clustal W\textsuperscript{39} in MEGA5\textsuperscript{37} was employed to align the amino acid sequence of the cloned genes with their reference sequences selected from the NCBI GenBank database. The evolutionary distances were calculated using a Poisson distribution model and phylogenetic trees were constructed with the Neighbor Joining method. 2000 bootstrap iterations were carried out to test the statistical significance of the phylogenetic trees\textsuperscript{37}.

**Plant inoculation**

After transformation of \textit{Agrobacterium} cells (GV3101) with all the six constructs, cells were grown on LB (Luria-Bertani) agar plates with specific antibiotics (Rifamycin 100 µg/ml, Carbenicillin 100 µg/ml, Streptomycin 50 µg/ml and Spectinomycin 20 µg/ml) at 28°C for 60 hours. Colony PCR (initial denaturation at 94°C for 2 min; 34 cycles of incubation at 94°C for 30 s, at 50°C for 1 min and 72°C for 2 min; final extension at 72°C for 10 min) with the primers used for the previous gene cloning was performed to identify the correct \textit{Agrobacterium} clones. Subsequently, a single colony for each
construct was picked from the plates and grown in LB medium with the corresponding antibiotics to a visible density.

Wildtype *N. benthamiana* plants were grown at 26°C with a photoperiod of 16 hours light and 8 hours dark. The PA-GFP construct (to express the GFP transgene) and the RNAi-GFP construct (to induce RNA silencing) were infiltrated into whole leaves of wildtype *N. benthamiana* plants using *Agrobacterium* at 4 leaf stage\(^4\). At 4 days post inoculation (dpi), the left halves of GFP-silenced leaves were inoculated with *Agrobacterium* containing the cloned gene or HC-Pro, the positive control in this study, while the right halves were infiltrated with empty PLH7000 vector as a negative control. Young leaves were inoculated with a mixture of *Agrobacterium* culture (500 μl for each construct) containing the corresponding constructs by placing a syringe of 5 ml on the surface of the wildtype *N. benthamiana* leaves and gently pressing the plunge until the whole leaves were water-soaked. All four leaves of a single wildtype *N. benthamiana* plants were infiltrated\(^4\). For each individual construct, three plants were used for repeating the inoculation experiments.

**Imaging and collection of samples**

Visual detection of green fluorescence in the *N. benthamiana* plants infiltrated with distinct constructs was performed using ultraviolet (UV) illumination. At 4 dpi, photographs were taken with a digital high-resolution camera under UV light. The inoculated areas of the leaves were cut off using a razor blade, frozen in liquid nitrogen and stored in a low temperature freezer for future analyses.
Results

Gene cloning of OYCrCMA-Rep, ALCCMA-Rep and CLCuGB-βC1 in pJET1.2

The coding sequences of the Alpha-Rep genes in OYCrCMA and ALCCMA and the βC1 gene in CLCuGB were amplified by PCR. Distinct DNA bands of approximately 850 bp, 950 bp and 350 bp were detected, indicating the successful amplification of OYCrCMA-Rep, ALCCMA-Rep and CLCuGB-βC1 (Fig. 4). The concentration of the purified PCR products was 62, 21 and 54 ng/μl, respectively.

Figure 4. PCR amplification of OYCrCMA-Rep (a), ALCCMA-Rep (b) and CLCuGB-βC1 (c). Lane 1 and 2 contain PCR products of the genes to be cloned. M refers to the 1kb DNA ladder plus marker (GeneRuler™). Lane 3 is a negative PCR control.

These DNA fragments were first cloned into the intermediate pJET1.2 vector in *E. coli* cells. Colony PCR was performed to select positive *E. coli* colonies containing the recombinant pJET1.2 construct by amplifying the corresponding genes with the pJET1.2 sequencing primers. Expected DNA bands of around 850 bp were obtained for the clones 6-24, suggesting that these clones contained the pJET1.2-OYCrCMA-Rep recombinant construct (Fig. 5a). Clones 6-10 were used for subsequent culture of *E. coli* cells. For the
pJET1.2-ALCCMA-Rep recombinant constructs, clones 10, 12, 14, 17, 18, 19, 20, 22 and 24 showed positive DNA bands of about 950 bp (Fig. 5b). Clones 10, 12, 17, 19, 20 and 22 were selected for further analysis. The DNA band of 350 bp was present in all the pJET1.2-CLCuGB-βC1 samples, indicating that all the colonies tested by colony PCR contained the recombinant construct. Clones 2, 4, 5, 6, 7 and 8 were chosen for subsequent analysis (Fig. 5c).

The recombinant plasmid constructs were purified from E. coli cultures and analyzed by digestion with Ncol and Xbal. The pJET1.2-OYCrCMA-Rep recombinant construct was cut into three DNA fragments, which was consistent with the fact that, in addition to the restriction sites introduced into both ends of the gene, another Ncol recognition site is located in the OYCrCMA-Rep gene (Fig. 6a). The recombinant plasmid clone 10 with the highest concentration of 177 ng/μl was used for further cloning of OYCrCMA-Rep in PLH7000. For pJET1.2-ALCCMA-Rep and pJET1.2-CLCuGB-βC1 recombinant constructs, two distinct DNA bands were detected on the agarose gels (Fig. 6b and c). The DNA band with large size was the pJET1.2 vector, while the other band was the corresponding cloned gene. Clone 17 of the pJET1.2-ALCCMA-Rep recombinant plasmids (142 ng/μl) and
clone 8 of the pJET1.2-CLCuGB-βC1 recombinant plasmids (133 ng/μl) were selected for subsequent cloning of the genes in PLH7000.

Figure 6. Digestion of the recombinant constructs pJET1.2-OYCrCMA-Rep (a), pJET1.2-ALCCMA-Rep (b) and pJET1.2-CLCuGB-βC1 (c) with Ncol and Xbal. The gel pictures show digestion products of the recombinant pJET1.2 constructs. The numbers above the gels stand for the clone numbers corresponding to those in Figure 5. M1 refers to the 1 kb DNA ladder plus marker (GeneRuler™), and M2 represents the 100 bp DNA ladder plus marker (GeneRuler™).
Sequencing and bioinformatics analysis

The cloned DNA fragments in pJET1.2 were sequenced, and the identified nucleotide sequences were translated into amino acid sequences using MEGA5 (Table 3). The amino acid sequence of the cloned OYCrCMA-Rep gene was 99% identical to that of OYCrCMA-Rep (CBK25781), and the cloned ALCCMA-Rep gene showed the highest amino acid sequence identity (97%) to ALCCMA-Rep (CBX51429). OYCrCMA-Rep (CBK25781) and ALCCMA-Rep (CBX51429) were the cloning templates used in this study. The relatively low identity is due to several nonsynonymous point mutations. For the cloned ALCCMA-Rep gene, there is also an insertion of one nucleotide because of a mutation occurring at the end of the cloned gene and leading to frameshift and truncation of the protein sequence. The predicted protein encoded by the cloned CLCuGB-βC1 gene showed only limited amino acid identity to CLCuGB-βC1 (CAQ63434), which is due to some mistakes in the primer design strategy. These mistakes lead to frameshift from the fourth amino acid and a truncation of the protein, from which it is inferred that the cloned CLCuGB-βC1 gene will be totally non-functional.

Table 3. Amino acid sequences of the cloned genes. The amino acids in blue are the amino acid sequence resulted from the cloning strategy. The amino acids in red are mutations generated in the process of gene cloning in pJET1.2. The second amino acid in each sequence was artificially introduced and not used for homology analysis.

<table>
<thead>
<tr>
<th>Clone</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>OYCrCMA-Rep</td>
<td>MAPSIKSVCWCFTLNFTGEIPQDLDFSNKGVYAVWQHEKVNHDHLQGGFFQFKGRRSLLLQAKVFDGYPHLEVMRAPSAEQAKKYCEKPESRVSGPWVFGEVFVPSGSNKRKLEEELANSNDNEAEPQRYRRAMAKKMNKDSHQALDNALPFELKEWQERLSGLISSPA DDRTIIWVYGPTGEGKSFARYLGLNKSIIYLPGGKSNDDMMYMYCKSPKCNLVIDYPRCNKEYINYSFLEMIKNRTIYSYKYPVGFIDPTCNHVIVMANFLPDYERISEDRIKLIDLS</td>
</tr>
<tr>
<td>ALCCMA-Rep</td>
<td>MAPSVQSSWWWCFTVFLSSTAPDLVPLFENTAVSYACWQEEESTPTTKRRLQGQLQCKGKRTLSEQKYLFGGLNPHLEKQRARKT DEARDYCMKEETRVSQGPFEGYVAAGSIIKRREAVERSPV RMSEENPSVRFRVKAGIAEAEQQSAPEIQIISNLKSWQLRLKQL SRRDPDDRTIFWVYGPHTGEGKSTFARLYRSGWVFRGGSADNVAYQYYIYGLGNNIVFDIPRDKITYLQYSLVEMFKDRLVSNKYEPIAMPLNClIHVVMNSFLPDMEKISSRDHVIPCISLWCLS</td>
</tr>
<tr>
<td>CLCuGB-βC1</td>
<td>MATINSGAPRESTSESMSSSKTIWTAR</td>
</tr>
</tbody>
</table>
Figure 7. A phylogenetic tree of Alpha-Rep constructed by neighbor-joining based on alignment of amino acid sequences. The sequences marked with a yellow asterisk were used to clone Rep genes and generate constructs used in this study, while sequences with a red asterisk were Rep gene clones constructed in pJET1.2 used for subsequent
PTGS suppression assays. The number of amino acid substitutions per site is in proportion to the length of the horizontal lines. Numbers on branches indicate bootstrap values. The accession numbers of the Alpha-Rep proteins are included in the sequence names. Alphasatellites used in this phylogenetic tree were *Gossypium darwinii* symptomless alphasatellite (GDarSLA), *Gossypium mustelinium* symptomless alphasatellite (GMusSLA), Guar leaf curl alphasatellite (GuLCuA), Hollyhock yellow vein alphasatellite (HoYVA), Cotton leaf curl Burewala alphasatellite (CLCuBuA), Melon chlorotic mosaic alphasatellite (MeCMA), Cotton leaf curl Gezira alphasatellite (CLCuGB), Papaya leaf curl alphasatellite (PapLCuA), Croton yellow vein mosaic alphasatellite (CroYVMA), Vernonia yellow vein Fujian alphasatellite (VYFvA), Chilli leaf curl Multan alphasatellite (ChLCuMuA), Okra yellow crinkle Cameroon alphasatellite (OYCrCMA), Tomato yellow leaf curl China alphasatellite (TYLCCNA), Malvastrum yellow mosaic Cameroon alphasatellite (MalYMCMa), Tobacco leaf curl Pusa alphasatellite (TbLCuPuA), Okra leaf curl alphasatellite (OLCuA) and Ageratum leaf curl Cameroon alphasatellite (ALCCMA).

A phylogenetic tree was constructed for Alpha-Rep (Fig. 7) using MEGA5. The analysis showed that the Alpha-Rep protein encoded by Melon chlorotic mosaic alphasatellite (MeCMA), an atypical alphasatellite associated with a bipartite begomovirus (Melon chlorotic mosaic virus; MeCMV) in the NW, grouped separately from all the other Alpha-Rep proteins. The remaining Alpha-Rep proteins encoded by alphasatellites associated with OW begomoviruses were phylogenetically divided into two major groups, the new alphasatellite group and the old alphasatellite group. The OYCrCMA-Rep sequence used in this study, other OYCrCMA-Rep sequences and the Rep protein encoded by Hollyhock yellow vein alphasatellite (HoYVA) grouped into one clade (100% bootstrap value). The OYCrCMA-Rep protein consisting of 289 amino acids grouped into the clade of Alpha-Rep of new alphasatellites and was evolutionarily distant from all the other Alpha-Rep proteins encoded by OW begomovirus associated alphasatellites. ALCCMA-Rep consisting of 315 amino acids, typical of Alpha-Rep proteins, grouped with the Alpha-Rep proteins encoded by Cotton leaf curl Gezira alphasatellite, Okra leaf curl alphasatellite and Malvastrum yellow mosaic Cameroon alphasatellite from Africa in the old alphasatellite major group (84% bootstrap value).

Phylogenetic analysis of βC1 (Fig. 8) also divided the βC1 proteins into two major groups, βC1 proteins encoded by betasatellites from Asia and those encoded by betasatellites from Africa. The CLCuGB-βC1 clone used grouped together with the other CLCuGB-βC1 sequences and the βC1 proteins encoded by another betasatellite from Africa, Ageratum leaf curl Cameroon betasatellite (95% bootstrap value).
Figure 8. A phylogenetic tree of βC1 constructed by neighbor-joining based on alignment of amino acid sequences. The sequence marked with a yellow asterisk was used to clone the βC1 gene and generate the construct used in this study. The number of amino acid
substitution per site is in proportion to the length of the horizontal lines. Numbers on branches indicate bootstrap values. The accession numbers of the βC1 proteins are included in the sequence names. Abbreviation of the betasatellites is according to Reference 42.

**Further cloning of the Alpha-Rep genes in PLH7000**

The binary PLH7000 vector and the pJET1.2-ALCCMA-Rep recombinant construct were digested with Ncol and Xbal (Fig. 9a). Digested PLH7000 vector and cloned DNA fragments were separated by electrophoresis, purified from the agarose gel and ligated together. In addition to complete digestion with Xbal, the construct pJET1.2-OYCrCMA-Rep was partially digested with Ncol, because a Ncol recognition site was located in the OYCrCMA-Rep gene and complete digestion with Ncol would cut the gene into two pieces (Fig. 9b). A DNA band of approximately 850 bp was expected for the complete OYCrCMA-Rep gene.

**Figure 9.** Digestion of the recombinant pJET1.2 constructs with Ncol and Xbal. a) Digestion of pJET1.2-ALCCMA-Rep. Lane 1 contained the PLH7000 vector after double digestion by both Ncol and Xbal. Lanes 2 and 3 contained the PLH7000 vector with single digestion by Ncol or Xbal, respectively. Lanes 4 and 5 contained digestion products of pJET1.2-ALCCMA-Rep. b) Partial digestion of pJET1.2-OYCrCMA-Rep. M1 refers to the 1 kb DNA ladder plus marker (GeneRuler™), and M2 represents 100 bp DNA ladder plus marker. Lane 1 contained the partial digestion products of pJET1.2-OYCrCMA-Rep. The DNA band marked with a red arrow was expected for the complete OYCrCMA-Rep gene.
After ligation, *E. coli* cells were transformed with the ligation products. The PLH7000 constructs were subsequently extracted from *E. coli* cells and tested by digestion with *Ncol* and *XbaI* to select the recombinant clones. The digestion pattern of the recombinant PLH7000 construct was similar to that of the recombinant pJET1.2 construct. On the gel picture, clone 4 showed two expected DNA bands, indicating that the construct in this sample contained the cloned OYCrCMA-Rep gene (Fig. 10a). As for the ALCCMA-Rep gene, a distinct DNA band of 950 bp was detected in the lanes for clones 3 and 5, suggesting that these two clones were recombinant constructs (Fig. 10b). Clone 5 with a DNA concentration of 246 ng/μl was selected for further transformation into *Agrobacterium* cells.

**Figure 10.** Digestion of the purified PLH7000 constructs with *Ncol* and *XbaI* to select the recombinant constructs. a) Digestion of the potential OYCrCMA-Rep recombinant constructs. b) Digestion of the potential ALCCMA-Rep recombinant constructs. Lanes 1-6 are digestion products of the PLH7000 constructs. M1 refers to the 1 kb DNA ladder plus marker (GeneRuler™), and M2 represents the 100 bp DNA ladder plus marker.

Colony PCR was used to select positive *Agrobacterium* colonies containing the recombinant PLH7000 constructs. For the PLH7000-OYCrCMA-Rep recombinant constructs, all the tested colonies showed positive bands of approximately 850 bp (Fig. 11a). DNA bands with the expected size were observed also for ALCCMA-Rep, but not as strong as those for OYCrCMA-Rep (Fig. 11b). These results showed that both of the recombinant PLH7000 constructs were successfully transferred into *Agrobacterium* cells.
RNA silencing suppression analysis

To assess the RNA silencing suppressor activity of the OYCrCMA-Rep and ALCCMA-Rep proteins, wildtype *N. benthamiana* plants were first inoculated with the PA-GFP construct to express the GFP transgene and the RNAi-GFP construct to induce transient GFP silencing. Silencing effect of GFP, characterized by appearance of red chlorophyll autofluorescence instead of green fluorescence, starts at 2 dpi. Initially, the phenomenon was observed at the veins, and then it spread to the whole inoculated leaves at 4 dpi. At this occasion, the left halves of the GFP-silenced leaves were inoculated with expression constructs encoding HC-Pro, OYCrCMA-Rep or ALCCMA-Rep, and the right halves were infiltrated with empty PLH7000 vector as a negative control. For each of the three constructs, fifteen young and fully expanded leaves of three wildtype *N. benthamiana* plants were used to perform RNA silencing suppression analysis, and each leaf was inoculated with 500 μl Agrobacterium culture containing the corresponding constructs. The tested leaves were observed under UV light to evaluate the RNA silencing suppressor activity of the Rep proteins. The leaves inoculated with OYCrCMA-Rep showed a much weaker green fluorescent signal than those infiltrated with HC-Pro, indicating that OYCrCMA-Rep is a relatively weak suppressor of RNA silencing. In contrast, a very strong fluorescent signal was detected on the leaves inoculated with ALCCMA-Rep. The suppression ability of ALCCMA-Rep was comparable to that of HC-Pro, which has been identified to be a strong suppressor of RNA silencing. It is possible to conclude that both OYCrCMA-Rep and ALCCMA-Rep have suppression activity of RNA silencing, and the suppressor activity of ALCCMA-Rep is much stronger than that of OYCrCMA-Rep.

Figure 11. Colony PCR for OYCrCMA-Rep (a) and ALCCMA-Rep (b) to select positive *Agrobacterium* colonies containing the recombinant PLH7000 constructs. M refers to the 1 kb DNA ladder plus marker (GeneRuler™). Lane 8 is a PCR negative control.
Figure 12. RNA silencing suppression analysis. The left halves of the GFP-silenced leaves were inoculated with expression constructs encoding HC-Pro (a), OYCrCMA-Rep (b) and ALCCMA-Rep (c), and the right halves were infiltrated with empty PLH7000 vector as a negative control. The yellow arrows indicate infiltration sites.
Discussion

RNA silencing is a natural defense mechanism of plants acting by siRNA to counteract viral attack. To overcome this defense mechanism, plant viruses and their associated satellites also develop their own mechanism acting through RNA silencing suppressors with the ability to repress silencing and promote systemic viral infection. Previous studies have identified a huge number of silencing suppressors encoded by viruses and their associated satellites. In fact, all classes of plant-infecting viruses including geminiviruses encode at least one RNA silencing suppressor.

Several proteins encoded by begomoviruses and their associated satellites have been identified to have suppression activity, but different viruses adopt different strategies to use these proteins as their weapons. One possible reason to explain this fact is the co-evolution of the RNA silencing mechanism of plants and different mechanisms adopted by plant viruses for suppressing RNA silencing. There is an arms race between them. Yeast two-hybrid assay and co-immunoprecipitation could be used to identify the targets of RNA silencing suppressors and dissect the mode of action of the suppression activity. Two bipartite begomoviruses closely related to each other inducing cassava mosaic disease in Africa, *East African cassava mosaic Cameroon virus* (EACMV) and ACMV, employ different proteins (TrAP and AC4, respectively) to suppress RNA silencing. Five gene products including C4, C2, V2 (from begomoviruses), βC1 (from betasatellites) and Alpha-Rep (from alphasatellites) can possibly become the suppressors in different begomovirus-satellite complexes to counteract RNA silencing. For the Alpha-Rep proteins, at least two Alpha-Rep proteins encoded by GDarSLA and GMusSLA have been shown to have suppression activity of RNA silencing. Alphasatellites were discovered one decade ago, but our knowledge about their function in the infection cycle of the begomovirus-betasatellite complexes is still very limited. The RNA silencing suppression activity of Alpha-Rep proteins could possibly explain the evolutionary advantage of begomovirus-alphasatellite complexes. In this study, the RNA silencing suppression activity of Alpha-Rep proteins encoded by two alphasatellites from Africa (OYCrCMA and ALCCMA) was investigated.

The results showed that both OYCrCMA-Rep and ALCCMA-Rep have suppression activity of RNA silencing, and the suppressor activity of ALCCMA-Rep is much stronger than that of OYCrCMA-Rep. Subsequent quantitative assays such as northern blots and reverse transcription real-time PCR (Q-RT-PCR) are required to further validate the results. Northern blot is an effective method to quantify the transcript level of individual genes. In this study, northern blots could be used to measure the mRNA of GFP. If a gene product has suppression activity of RNA silencing, the transcript level of GFP in the plant sample, in which this gene is expressed, is expected to be
increased. The procedure for northern blot analysis of GFP transcript levels is as follows. Total RNA is extracted with appropriate method from left and right halves of leaves (negative control), respectively. Through agarose gel electrophoresis, the mRNA molecules are separated according to size, transferred to a nitrocellulose or nylon membrane by electrolytic or capillary blotting and fixed to the membrane through ultraviolet cross-linking. DNA or RNA probes complementary to GFP mRNA are labeled with radioactivity or fluorescent dye. After hybridization of GFP transcripts with probes, fluorescent or radioactive signals, which are proportional to the quantity of GFP transcripts, can be detected. An internal control is necessary to normalize the results. Housekeeping genes, which are considered to have the same transcript level among samples with different treatment such as ribosomal RNA, are reprobed on the membrane and are used to perform normalization of the transcript level of GFP\textsuperscript{48-50}. There are several advantages using the method of northern blot. However, compared to other methods used to determine the transcript levels of specific genes, such as Q-RT-PCR, northern blot has low sensitivity. It is only semi-quantitative, which means that it can only differentiate great differences more than two fold among samples with distinct treatment\textsuperscript{47}. So, if a small difference of suppression activity among different gene products needs to be clarified, Q-RT-PCR is advisable. Another important test is to quantify siRNA levels for the GFP gene.

The phylogenetic analysis showed that the OYCrCMA-Rep protein consisting of 289 amino acids grouped into the clade of Alpha-Rep of new begomovirus alphasatellites and that it was evolutionarily distant from all the other Alpha-Rep proteins. ALCCMA-Rep consisting of 315 amino acids, typical of alphasatellites, grouped to Rep proteins of old begomovirus alphasatellites from Africa. A previous study has demonstrated that the Rep protein encoded by an old alphasatellite (GDarSLA) also showed stronger suppressor activity than that encoded by a new alphasatellite (GMusSLA)\textsuperscript{12}. It seems that Rep proteins of old begomovirus alphasatellites have higher suppression ability of RNA silencing than those of new alphasatellites. Additional studies on the RNA silencing suppression activity of Alpha-Rep is needed to support this hypothesis. Sequence analysis of Alpha-Rep proteins could be conducted to identify conserved and variable regions for Alpha-Rep sequences of new alphasatellites and those of old alphasatellites.

In the process of gene cloning, several nonsynonymous point mutations have been generated in both the cloned OYCrCMA-Rep gene and the cloned ALCCMA-Rep gene. For the cloned ALCCMA-Rep gene, there is also an insertion of one nucleotide because of a mutation occurring at the end of the cloned gene and leading to frameshift and truncation of the protein sequence. These mutations could exert positive or negative effects on the RNA silencing suppression activity of Alpha-Rep proteins.
References


RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J. Virol. 79, 2517–2527.


Acknowledgements

I sincerely thank my supervisors Anders Kvarnheden and Naeem for their great support on my project.

Also many thanks to all the very kind members of the plant virology group who helped me.

At last, I would like to thank my dear parents. Thank you for your care during the writing of my thesis.