Novel insertion sequence-like elements in phytoplasma strains of the aster yellows group are putative new members of the IS3 family

I.-M. Lee *, Y. Zhao, K.D. Bottner

Molecular Plant Pathology Laboratory, USDA, ARS, Beltsville, MD 20705, USA

Received 16 September 2004; received in revised form 2 November 2004; accepted 17 November 2004

First published online 8 December 2004

Edited by H.B. Deising

Abstract

Novel insertion sequence (IS)-like elements were isolated and characterized from phytoplasma strains in the aster yellows (AY) group (16SrI). The IS-like elements were cloned from phytoplasma strains AY1 and NJAY or PCR-amplified from 15 additional strains representing nine subgroups in the AY group using primers based on sequences of the putative transposases (Tpases). All IS-like elements contained sequences encoding similar Tpases of 321 amino acids (320 for strain CPh). Substantial amino acid sequence variability suggested multiple species of Tpases or IS-like elements exist in the AY phytoplasma group. These Tpases have an identical DDE motif that is most similar to the DDE consensus of Tpases in the IS3 family.

© 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Insertion sequence; Aster yellows phytoplasma group; Peanut witches'-broom phytoplasma group; Mexican periwinkle virescence phytoplasma group.

1. Introduction

Phytoplasmas are plant pathogenic Mollicutes that cannot be cultured in vitro. To date, several hundred strains representing 15 16S rDNA RFLP or more than 20 phylogenetic groups have been identified by molecular means [1–3]. The aster yellows (AY) phytoplasma group (16SrI) strains are associated with more than 100 economically important diseases worldwide and represents the most diverse group known [4–6]. Currently, the AY group is differentiated into 16 subgroups on the basis of RFLP analysis of 16S rRNA gene sequences [6]. Moreover, there is considerable genetic variation among strains in some subgroups. Earlier studies on the genomic organization revealed extensive chromosomal rearrangements among strains in a given subgroup, and the presence of multiple uncharacterized repetitive sequences in chromosomes of strains in the AY phytoplasma group [4]. Often, genetic variations among strains coincide with the pathogenic variations. The primary mechanism that is attributable to genetic variations among phytoplasmas is unknown.

Insertion sequences (ISs) are a group of small, mobile genetic elements. In prokaryotes, IS elements are known to effect chromosomal rearrangements and the expression of neighboring genes. Genomic mutation caused by IS elements is an important mechanism for generating genetic diversity. The presence of insertion sequence elements in the chromosomes of phytoplasmas had not been investigated and verified until the recent report of an IS-like element in peanut witches'-broom (PhWB) phytoplasma (GenBank accession no. AY270153).

* Corresponding author. Tel.: +1 301 504 6024; fax: +1 301 504 5449.
E-mail address: leeim@ba.ars.usda.gov (I.-M. Lee).
Typical IS elements are between 800 and 2500 bp long and contain two important components, a gene encoding a transposase (Tpase) (typically 250–400 amino acids long) and inverted repeat (IR) structures (usually 10–40 bp long) at the termini. In prokaryotes, more than 500 such elements have been identified and classified into at least 17 families on the basis of one or more of the following criteria: (i) IS open reading frame (ORF) organization, (ii) conserved signature DDE motifs, an essential part of the catalytic domain among Tpases, (iii) similarity of IR sequences, and (iv) length of target site duplications [7]. In the present study, we report the discovery of several distinct but closely related IS-like elements in the AY group phytoplasmas. These IS-like elements encode Tpases that share conserved signature motifs and sequence similarity with those of the IS3 family [7]. A preliminary synoptic report has been published in abstract form [8].

2. Materials and methods

2.1. Phytoplasma strains and DNA extraction

Twenty three phytoplasma strains representing nine subgroups in the aster yellows phytoplasma group (16Sr I) [9] and strain Mexican periwinkle virescence (MPV), belonging to group 16Sr XIII [2], were used in this study (Table 1). Two symptom type variants of strain IA4 were also included in this study. Strain variant IA4Ph exhibited typical phyllody (the development of floral parts into green leafy structures) or virescence (the development of green flowers) symptoms, while strain variant IA4W exhibited small normal flower structure with normal pigment. Total nucleic acid was extracted from phytoplasma infected tissues according to previously described methods [5,6].

2.2. Cloning and sequencing of IS-like elements

Plasmid probe pAY27, previously constructed based on a cloned DNA fragment from AY phytoplasma strain AY1 (subgroup 16SrI-B), was shown to hybridize to multiple regions of genomic DNA from AY phytoplasma strains [4]. This probe was sequenced and a BLAST search of the NCBI protein database was performed to search for the presence of a putative IS. Primers used in this study for PCR amplification and sequencing are as follows: AYIS-F0 (5′-gcggtcatcgtaaaatcactga-3′), AYIS-F01A (5′-ataaaaatcccccagttgcgc-3′), AYIS-F01 (5′-ccaacaaaatcgcattaaagc-3′), AYIS(rt)F1 (5′-acaaccaaattatcgcttccc-3′), AYIS-F3 (5′-gggtacttaaatata-tagtgtaagg-3′), AYIS-R1 (5′-gacgttatc-
acgtggatgc-3'), AYIS-R4 (5'-gctaaatcattgattgttcc-3'), and AYIS-R5A (5'-ggegtaaaaaagtttaaagtt-3').

Partial genomic libraries were constructed from strains AY1 and NJAY (subgroup 16SrI-A). EcoRI-digests of genomic DNA were cloned into Escherichia coli strain TOP10 (Invitrogen, Carlsbad, CA) using the cloning vector pUC19 according to standard methods [10]. The primer pair AYIS-F0/AYIS-R1 was designed on the basis of the putative IS detected in plasmid pAY27 and used to screen for IS-containing clones from the two genomic libraries. The nucleotide sequences of putative IS elements in selected IS-containing clones were determined by the dideoxy method using an autosequencer (Applied Biosystems PRISM model 3100). The identities of putative IS elements were determined by BlastP searches against the LMGM (available in GenBank) were aligned by using the Clustal W algorithm [10]. The bootstrap neighbor-join tree was constructed using phylogenetic analysis program Clustal X (Version 1.8) [11], and DNASTAR’s Laser Gene software (DNASTAR, Madison, WI, USA). Cladistic analyses of phytoplasma sequences were performed with PAUP (phylogenetic analysis using parsimony), version 4.0 written by D.L. Swofford (University of Illinois), on a Power Mac G4.

2.4. Sequence and phylogenetic analyses of IS-like elements

Phylogenetic analyses were performed based on nucleotide or deduced amino acid sequences of Tpases. The Tpase sequences of the IS3 family are concatenated amino acid sequences that combined peptides encoded by open reading frames, orfA and orfB, of the Tpase gene. Sequences of Tpases in AY group phytoplasma strains and in representative members of the IS3 family (available in GenBank) were aligned by using the Clustal W algorithm [10]. The bootstrap neighbor-join tree was constructed using phylogenetic analysis program Clustal X (Version 1.8) [11], and DNASTAR’s Laser Gene software (DNASTAR, Madison, WI, USA). Cladistic analyses of phytoplasma sequences were performed with PAUP (phylogenetic analysis using parsimony), version 4.0 written by D.L. Swofford (University of Illinois), on a Power Mac G4.

3. Results and discussion

3.1. Detection of putative IS elements in diverse AY group phytoplasmas

A BLAST search revealed that one of the ORFs in the cloned DNA fragment pAY27 encodes a putative Tpase that shares high homology (50–69% identity) with uncharacterized Tpases associated with PnWB phytoplasma (group 16SrII) and onion yellows phytoplasma (GenBank accession no. AP006628), and shares a significant homology (about 36% identity and 55% similarity) with IS elements from low G + C gram-positive bacterial species of the IS3 family. We have identified several clones containing putative IS elements using partial genomic libraries constructed from strains AY1 and NJAY. Six IS-containing clones (two from strain AY1 and four from strain NJAY) were sequenced and analyzed. Each of the six clones contained one ORF encoding a putative Tpase consisting of 321 amino acids. The resulting six IS-like elements were designated AY1-IS65, AY1-IS82, NJAY-IS40, NJAY-IS43, NJAY-IS48, and NJAY-IS94. Nucleotide and deduced amino acid sequences of the IS-like element NJAY-IS48 is shown in Fig. 1.

The PCR assay using Tpase-specific primers revealed that an additional 15 phytoplasma strains representing nine subgroups in the AY group (Table 1) and a strain...
belonging to the Mexican periwinkle virescence (MPV) group (16SrXIII), contain IS-like elements. The PCRs using primer pair AYIS-F01A/AYIS-R5A generated a DNA fragment (about 1.1 kbp) from each of nine AY phytoplasma strains (AV2192, AVUT, BB, BBS-3, CPh, STRAWB2, PRIVC, PVM, and HyPH) representing seven 16Sr group or subgroup affiliations, which contains the entire sequence encoding a putative Tpase gene. We also

---

### IS3 consensus

<table>
<thead>
<tr>
<th>IS3 consensus</th>
<th>tDiTy (58-60)</th>
<th>HsDqGsu-y-s (35)</th>
<th>s-G----dN----Esf---lK</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJAY43 (16SrI-A)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>NJAY48 (16SrI-A)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>NYDP (16SrI-A)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>PVM (16SrI-A)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>AY1-65 (16SrI-B)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>AY1-82 (16SrI-B)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>HyPH (16SrI-B)</td>
<td>tYlTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>PRIVC (16SrI-B)</td>
<td>tYlTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>OY1 (16SrI-B)</td>
<td>tDiTc (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>OY2 (16SrI-B)</td>
<td>tDiTc (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>OY5 (16SrI-B)</td>
<td>tDiTc (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>OY9 (16SrI-B)</td>
<td>tDiTc (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>shkanprdNaviEnffgqmK</td>
</tr>
<tr>
<td>CPh (16SrI-C)</td>
<td>tDiTc (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatpldNaviEnffgqmK</td>
</tr>
<tr>
<td>STRAWB2 (16SrI-K)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>AV2192 (16SrI-L)</td>
<td>tDiTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatprdNaviEnffgqmK</td>
</tr>
<tr>
<td>AVUT (16SrI-L)</td>
<td>tYlTy (56)</td>
<td>HsDqGvuqyq (35)</td>
<td>srkatpldNaviEnffgqmK</td>
</tr>
</tbody>
</table>

---

Fig. 2. Alignment of the extended regions of the DDE motif of IS transposons from AY phytoplasmas and the consensus sequence of the IS3 family. The amino acid residues shown in bold uppercasse constitute the key part of the conserved motif. The amino acid residues shown in lowercase are those predominant within the IS3 family. The numbers in parentheses indicate the number of amino acid residues between the conserved acidic amino acids. 16Sr group or subgroup affiliations are listed in parentheses. The IS3 family consensus was retrieved from the LMGM-CNRS’ Insertion Sequence Database at http://www.is-biotoul.fr/is.html. Asterisks indicate the three acidic amino acids. OY1, OY2, OY5, and OY9 transposases are based on the OY-M sequence available in GenBank (accession no. AP006628) and the PnWB transposase is based on the PnWB sequence available in GenBank (accession no. AT207153).
detected IS-like elements in strains ACLR-AY and PaWB, representing two additional subgroups, by PCR using primer pair AYIS-F0/AYIS(r)R1 which generated a DNA fragment containing a portion (about 400 bp) of the putative Tpase gene sequence. DNA fragments containing putative Tpase sequences were also generated from strain MPV using both primer pairs. The IS-like element designations and GenBank accession numbers are as follows: AY1-IS65, AY497459; AY1-IS82, AY364444; NJAY-IS40, AY497463; NJAY-IS43, AY497460; NJAY-IS48, AY497461; NJAY-IS94, AY497462; ACLR(AY)-IS, AY497464; BBS3-IS, AY497465; CHRY-IS, AY497466; MBS-IS, AY497467; PaWB-IS, AY497468; CPh-IS, AY497469; AV2192-IS1, AY739440; AVUT-IS1, AY739441; AVUT-IS2, AY739442; BB-IS1, AY739443; BBS3-IS1, AY739444; CPh-IS1, AY739445; MPV-IS1, AY739446; STRAWB2-IS1, AY739447; PRIVC-IS1, AY739448; PVM-IS2, AY739449; HyPH-IS2, AY739450. The results of sequencing indicated that not all IS-like elements encode a functional Tpase. NJAY-IS40, NJAY-IS94, MBS-IS, ACLR (AY)-IS, PaWB-IS, and MPV-IS contain degenerate Tpase nucleotide sequences. All the functional Tpases contained 321 amino acids, except strain CPh which contained 320 amino acids.

3.2. Characterization and identification of AY IS-like elements

In the present study, all putative Tpases determined in AY group phytoplasma strains, as well as the one encoded by the peanut witches'-broom (PnWB) phytoplasma-associates IS-like element, have a conserved DDE motif that is most similar to the DDE consensus of the Tpases of the IS3 family [7]. Each IS-like element contains internal G + C rich blocks, a characteristic of the IS3 family. However, many IS-like elements identified in the AY phytoplasma group in this study lack apparent IR sequences at both ends of the Tpase. The majority of IS-like elements in the AY phytoplasma group contain only one ORF in Tpases. Phylogenetic analysis based on the deduced amino acid sequences of Tpases revealed that IS-like elements in the AY phytoplasma group and two from phytoplasmas belonging to PnWB and MPV phytoplasma groups constituted a distinct cluster within the IS3 family (Fig. 3).

3.3. Genetic heterogeneity of AY IS-like elements

Based on deduced amino acid sequences of encoded Tpases, AY IS-like elements from various sources are genetically heterogeneous. The two IS-like elements associated with strain NJAY, NJAY-IS43 and NJAY-IS48, are most closely related to each other, encoding identical Tpases (with 100% identities in amino acid sequences). IS-like elements associated with other strains in the AY group phytoplasmas encode heterogeneous
Tpases with substantial sequence variations (nucleotide or deduced amino acid sequences), ranging from 9.3% to >30%. Tpases among AY phytoplasmas in North America share high sequence homologies (88–100%). However, they are significantly different from those associated with onion yellows (OY) phytoplasma in Japan, sharing only 42–78% sequence homology. Moreover, unlike Tpases in North American AY phytoplasma strains, Tpases in OY phytoplasma consist of only 309–316 amino acids. The AY IS-like elements isolated from various sources seem to consist of multiple species.

Phylogenetic analyses based on nucleotide (Fig. 4(a)) or deduced amino acid (Fig. 4(b)) sequences of the Tpases also indicated that there were several distinct lineages of IS-like elements in AY group phytoplasma strains. These various IS-like elements were delineated into two distinct phylogenetic clusters: (i) IS-like elements in AY phytoplasma strains from North America and Europe, and (ii) IS-like elements in onion yellows and...
peanut witches'-broom (belonging to 16SrII) phytoplasmas of Asian origin.

3.4. Presence of AY-IS elements in genomes of AY group phytoplasmas

Subgroups 16SrI-A and 16SrI-B consist of diverse phytoplasma strains that are associated with various plant hosts [6]. The distribution of IS-like elements in chromosomes of several representative strains (subgroup 16SrI-A, strains NJAY, OKAY1Ph, IA4Ph, IA4W, OnionD3, OnionD8, and BB; 16SrI-B, strains AY1, SAY, OKAY9B, and ErY) (Table 1) from various sources was determined by Southern blot hybridization analysis of EcoRI-digested genomic DNA. The results revealed multiple copies of IS-like elements present in all strains tested. Hybridization profiles differed between strains of each subgroup as well as among strains in a given subgroup (Fig. 5(a)). The profiles are more diverse in 16SrI-B strains than in 16SrI-A strains. In many cases, the profiles vary with strains from different sources. The variations seem to be consistent with the much wider plant host range and diverse insect vectors of the 16SrI-B strains in nature. The two symptom-type variants (phyllody and non-phyllody) of strain IA4 also exhibited different hybridization profiles.

Southern hybridization analyses of uncut genomic DNA from AY group phytoplasmas revealed that the probe hybridized to both chromosomal and putative extrachromosomal DNA fragments (Fig. 5(b)). The probe hybridized to chromosomal bands (above 21.2 kb) in all five AY phytoplasma stains (NJAY, CN1, AY1, SAY and CPh) but not in the sweet potato witches'-broom (SPWB) phytoplasma strain (16SrII-A). This probe also hybridized to putative extrachromosomal bands (at positions slightly below 21.2 kb and/or between 3.0 and 5.1 kb) in four AY phytoplasma strains (NJAY, CN1, AY1, and SAY). This implies that IS-like elements may also be present in plasmids.

In summary, we have identified novel IS-like elements in various strains of the AY phytoplasma group. These AY IS-like elements are closely related to one another and form a distinct cluster. The conserved signature motifs and sequence similarity of encoded Tpases and phylogenetic analysis based on the deduced amino acid sequences of the Tpases suggests that the IS-like elements in the AY group phytoplasmas may represent a new group or subfamily in the IS3 family. However, other features such as an ORF shift and characteristic terminal IR structures that are typical of the IS3 family seem to be lacking in the putative IS-like elements analyzed.

IS-like elements are widely present in the AY group phytoplasmas representing at least nine 16SrI subgroups. Substantial genetic heterogeneities are evident among these elements in various strains from different sources. Based on sequence dissimilarity and phylogenetic analyses of encoded Tpases, IS-like elements in the AY group phytoplasmas may represent several distinct species. Interestingly, closely related IS-like elements are also present in PnWB (16SrII) and MPV (16SrXIII) phytoplasmas, implying that horizontal transfers may have occurred within the AY group phytoplasmas and between phytoplasmas in the three phytoplasma groups. The evidence that the IS-like elements may be encoded by plasmids reinforces this notion.

IS elements are known to generate chromosomal rearrangements and to effect expression of neighboring genes. Transposition activity of IS elements results in increasing genetic diversity of bacterial pathogens. Southern hybridization studies in the present work indicated that members of AY subgroups 16SrI-A and 16SrI-B have undergone extensive chromosomal rearrangements with each strain exhibiting different hybridization profiles. Preliminary results suggest that transposition activity of IS-like elements in strain IA4 may be involved in the generation of the two symptom-type variants.

Acknowledgements

We thank C. Marcone, E. Seemu¨ller, and J. Fletcher for providing phytoplasma DNA or infected plant tissue.

References


