Complete nucleotide sequence of sugarcane streak
Monogeminivirus

F. L. Hughes¹, E. P. Rybicki¹, and R. Kirby²
¹Department of Microbiology, University of Cape Town, Rondebosch, and
²Department of Microbiology, Rhodes University, Grahamstown, South Africa

Accepted February 21, 1993

Summary. The complete nucleotide sequence of the genome of the Monogeminivirus sugarcane streak virus (SSV) was determined from cloned replicative form DNA. The genome is contained in one DNA circle of 2 758 nucleotides, and has four open reading frames with the potential to encode proteins of MW > 10 kDa: two in the viral (+) sense and two in the complementary (−) sense. Each open reading frame has a counterpart among the open reading frames reported for other Monogeminiviruses. A potential binding site for a DNA replication primer and potential transcriptional control sequences were identified on the (+) strand, and a possible intron on the (−) strand. Phylogenetic analysis of coat protein and replication-associated protein sequences of SSV and other grass-infecting geminiviruses indicate that SSV, although distinct from any other virus, is part of an “African streak virus subgroup” of Monogeminiviruses.

Introduction

Members of the taxonomic family Geminiviridae are plant viruses with geminate particles encapsidating single-stranded circular DNA. The genus Monogeminivirus — formerly Subgroup I — includes viruses with monopartite genomes which are obligately leafhopper-transmitted and mainly infect Gramineae (ICTV proceedings 1992; G. P. Martelli, pers. comm.). Maize streak virus (MSV) is the type member; other sequenced examples are the monocot-infecting wheat dwarf virus (WDV), digitaria streak virus (DSV), chloris striate mosaic virus (CSMV), miscanthus streak virus (MiSV), panicum streak virus (PanSV) [5, 21], and the dicot-infecting tobacco yellow dwarf virus (TobYDV; [28]).

The causative agents of streak disease of sugarcane (Saccharum officinarum) in Africa and Mauritius are known to be geminiviruses; these have been assumed to be strains of MSV on the basis that they infect maize, are transmitted by
the same vector as MSV (the leafhopper *Cicadulina mbila*), and are antigenically related (albeit distantly) to MSV [7, 32, 33]. However, the genomic DNAs of the sugarcane viruses cross-hybridize only weakly with one another, and not at all with MSV DNA, and the restriction endonuclease maps are very different from those of MSV isolates, and from one another [19, 20]. Therefore, it was previously proposed that the Natal sugarcane virus is a distinct virus (sugarcane streak virus, SSV) rather than a strain of MSV, and that the Natal and Mauritius sugarcane viruses may well also be distinct [19, 36].

In this paper we present the complete nucleotide sequence of SSV-Natal and compare it at both the nucleotide and derived amino acid level with other sequenced Monogeminiviruses. The phylogenetic relationship of SSV to other Monogeminiviruses is deduced, and comparative inferences are drawn on determinants of transcriptional specificity and host specificity in cereal geminiviruses.

**Materials and methods**

*Source of sugarcane streak virus DNA*

Streak-diseased sugarcane (*Saccharum officinarum*) cv. Uba was obtained from K. Harborne, South African Sugar Association Experimental Station (SASAES, Mt. Edgecombe, Natal, South Africa), and maintained by vegetative propagation in an insect-free plant room [19].

*DNA cloning and sequencing*

Total plant DNA containing double-stranded replicative form (RF-) DNA of SSV was extracted as described by [19]. Viral DNA was cloned in both orientations into the plasmid vector Bluescript SK (Stratagene, California) as a near full-length *Pst*I fragment, and as two separate *Bgl*II fragments.

Sets of ordered deletions were constructed for the *Pst*I clones, and for the clones containing the larger of the two *Bgl*II fragments [15]. DNA was sequenced by the deoxynucleotide chain termination method [39], using the Sequenase II kit (United States Biochemical Corporation), pUC sequencing primers (Bethesda Res. Labs.), and 35S-dATP (400 Ci/mmol) (Amersham). Sequence ambiguities due to band compressions were eliminated by the addition of dimethylsulphoxide (10% final concentration) into the labeling and termination mixes, which were incubated at 15°C and 45°C, respectively. The smaller *Bgl*II fragment was sequenced entirely in both directions. In order to sequence across a gap in the sequence of the virion strand, a 17-mer oligonucleotide primer (5'-TTGAGCGACGGCTAGG-3') was purchased (Beckman Instruments, Cape Town).

*Sequence analysis*

SSV sequence information was stored, assembled and analysed using GENEPRO Ver. 4.20 software (Riverside Scientific, Seattle), and the Genetics Computer Group, Inc. (GCG) VAX mainframe package, Versions 6.1 and 7.1 [8]. The sequences of MSV-N (Nigerian isolate, [30]), MSV-K (Kenyan isolate, [18]), MSV-S (South African isolate, [23]), CSMV ([2], WDV [25], DSV [9], and tobacco yellow dwarf virus (TobYDV, [28]) were obtained from the GENBANK or DDBJ databases. Sequences of MiSV [6] and PanSV [5] were obtained from the authors.