Sequence Diversity and Virulence in *Zea mays* of *Maize Streak Virus* Isolates

D. P. Martin,* J. A. Willment,* R. Billharz,* R. Velders,* B. Odhiambo,† J. Njuguna,‡ D. James,* and E. P. Rybicki*,1

*Department of Moleculare Cell Biology, University of Cape Town, Private Bag, Cape Town, South Africa, 7701; †Kenya Agricultural Research
Institute, Kabete, P.O. Box 14733, Nairobi, Kenya; and ‡Kenya Agricultural Research Institute, Muguga, P.O. Box 30148, Nairobi, Kenya

Full genomic sequences were determined for 12 *Maize streak virus* (MSV) isolates obtained from *Zea mays* and wild grass species. These and 10 other publicly available full-length sequences were used to classify a total of 66 additional MSV isolates that had been characterized by PCR–restriction fragment length polymorphism and/or partial nucleotide sequence analysis. A description is given of the host and geographical distribution of the MSV strain and subtype groupings identified. The relationship between the genotypes of 21 fully sequenced virus isolates and their virulence in differentially MSV-resistant *Z. mays* genotypes was examined. Within the only MSV strain grouping that produced severe symptoms in maize, highly virulent and widely distributed genotypes were identified that are likely to pose the most serious threat to maize production in Africa. Evidence is presented that certain of the isolates investigated may be the products of either intra- or interspecific recombination.   © 2001 Academic Press

Key Words: geminivirus; mastrevirus; recombinant genomes; recombination detection software; evolution; epidemiology.

INTRODUCTION

Maize streak virus (MSV) is the causal agent of maize streak disease, the most important disease of maize in Africa (Bosque-Perez, 2000). It is the type member of the genus *Mastrevirus* (family: *Geminiviridae*) and is relatively closely related to the less economically significant African and Australasian streak viruses *Panicum streak virus* (PanSV), *Sugarcane streak virus* (SSV), *Sugarcane streak Réunion virus* (SSRV), sugarcane streak Egypt virus (SSEV), and *Digitaria streak virus* (DSV) (Bigarré *et al.,* 1999; Briddon *et al.,* 1992; Briddon, 1996; Donson *et al.,* 1987; Hughes *et al.,* 1992). Together with other members of the genus these viruses have a single genomic component and an almost identical arrangement of genes (Rybicki *et al.,* 2000).

Since 1984 the full genomic sequences of nine unique MSV isolates have been published. Until recently all sequenced isolates had been obtained from severely symptomatic maize and shared greater than 95% sequence identity (Peterschmitt *et al.,* 1996). Full genomic sequences have now also been determined for three MSV isolates obtained from wild annual grass species and wheat (Schnippenkoetter *et al.,* 2001; Willment, 1999). These sequences share less than 90% identity with those of the isolates obtained from maize.

Various studies using partial nucleotide sequencing of MSV genomes and/or restriction fragment length polymorphism (RFLP) analysis of MSV-derived PCR products

have attempted to determine both the extent of MSV diversity and the distribution of MSV genotypes in Africa (Briddon *et al.,* 1994; Rybicki *et al.,* 1998; Willment *et al.,* 2001). Collectively these studies have involved the analysis of only 28 isolates from maize and 11 isolates from annual grasses and wheat. These investigations have confirmed that while virtually all isolates from maize shared greater than 95% nucleotide sequence identity (Briddon *et al.,* 1994; Rybicki *et al.,* 1998; Willment, 1999), isolates from wheat and annual grasses were considerably more diverse and shared between ~89 and ~78% nucleotide sequence identity with the maize isolates (Rybicki *et al.,* 1998; Willment *et al.,* 2001). All of these studies, however, concentrated on inferring phylogenies of isolates from nucleotide sequence data obtained from less than half of the isolates' genomes, and none fully explored the relationship between isolate genotypes and virulence in maize.

Here we describe the PCR–RFLP typing of 49 MSV isolates and the complete sequencing and analysis of 12 new MSV genomes selected from among the 49 as being representative of major virus groupings. These and 10 other publicly available full-length MSV sequences were used to classify a total of 85 MSV isolates into tentative strain and subtype groupings for which geographical and host distributions were determined. Agroinoculation of differentially resistant maize genotypes with 18 agroinfectious MSV isolates and symptom quantification by image analysis were used to identify specific groups of MSV genotypes that are likely to pose serious threats to maize production. Recombination between viruses was assessed using custom-written analysis software, and

the potential contribution of less virulent virus strains and species to the evolution of economically significant MSV genotypes was investigated.

RESULTS

Full genome sequencing and classification of MSV isolates

To enable the classification of MSV isolates characterized in this and other studies using PCR–RFLP and partial sequence analysis, isolates that represented as much as possible of the currently observed MSV diversity were selected for full genome sequencing. Forty-nine MSV isolates were analyzed using the PCR–RFLP technique developed by Willment *et al.* (2001). The isolates were of four major types: a "maize type" (represented by 37 isolates) found almost exclusively in maize, a "grass type" (represented by 10 isolates) found in a range of predominantly wild grass species, a "*Urochloa* type" (represented by MSV-Raw), and a "*Digitaria* type" (represented by MSV-Pat).

To determine the actual nucleotide sequence relationships between viruses in the different groups, eight maize-type isolates (MSV-MatA, MSV-MatB, MSV-MatC, MSV-Sag, MSV-Ama, MSV-Gat, MSV-MtKA, and MSV-MakD), two grass-type isolates (MSV-Mom and MSV-Jam), MSV-Raw, and MSV-Pat were cloned and fully sequenced (GenBank Accession Nos. AF329878–AF329889). A pairwise nucleotide sequence identity matrix was constructed using these 12 sequences and 10 other full-length MSV sequences currently available in GenBank (the matrix can be found on our Web site: ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/figa.pdf). To classify the sequenced isolates, a strain was arbitrarily defined as a group of isolates sharing greater than 94% nucleotide sequence identity. Based on this definition, the 22 sequences assessed contain representatives of five MSV strains. We have tentatively named these MSV-A, -B, -C, -D, and -E (Fig. 1).

To further classify the large group of MSV-A isolates, a subtype was arbitrarily defined as a group of isolates sharing greater than 98% sequence identity. By this definition the MSV-A sequences currently available could be divided into six subtypes, all of which were supported by phylogenetic analysis of the sequences. These were tentatively named MSV-A_1, -A_2 -A_3, -A_4, -A_5 and -A_6 (Fig. 1).

Geographical and host distribution of MSV strains and subtypes

We assigned strain and subtype classifications to PCR–RFLP typed isolates from both this study and that carried out by Willment *et al.* (2001) by comparing their RFLPs to those of the fully sequenced isolates. Similarly, isolates previously typed by partial nucleotide sequence analysis (Briddon *et al.,* 1994; Rybicki *et al.,* 1998) were
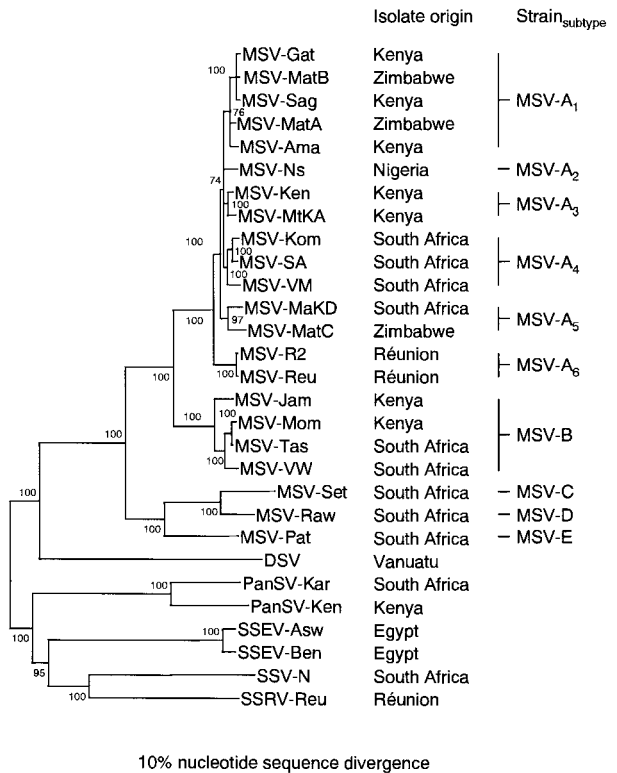

FIG. 1. A rooted neighbor joining tree (Saito and Nei, 1987) indicating the relationships between the MSV genomes sequenced in this study and all other publicly accessible full-length African streak virus genomic sequences. GenBank accession numbers for the sequences not determined in this study can be found on our Web site (ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/tableb.pdf). The MSV sequences can be divided into five major strain groupings (designated A–E). The MSV-A isolates were subdivided into six subtypes for purely descriptive purposes. The tree has been rooted on the sequence of the mastrevirus *Chloris striate mosaic virus* (not shown). Numbers associated with nodes represent the percentage of 1000 bootstrap iterations supporting the nodes. Nodes with less than 60% bootstrap support were collapsed.

also classified based on their relatedness to the fully sequenced isolates. Collectively, 85 MSV isolates were typed and used to assess both the prevalence of different MSV strains in nonmaize hosts and the distribution of different MSV-A subtypes in maize throughout Africa (Fig. 2). A listing of the strain and subtype classifications of each of the 85 MSV isolates can be found on our Web site: ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/tablea.pdf.

MSV-B was both the only detectable strain found in cultivated species, such as wheat and rye, and the most commonly detected strain in wild grasses (70%; Fig. 2). Although strain B viruses were also detected in maize, they represented only 3% of all isolates detected in this host. Whereas MSV-A subtypes accounted for 20% of the isolates detected among wild grasses (none was from wheat or rye), strain C, D, and E viruses are apparently
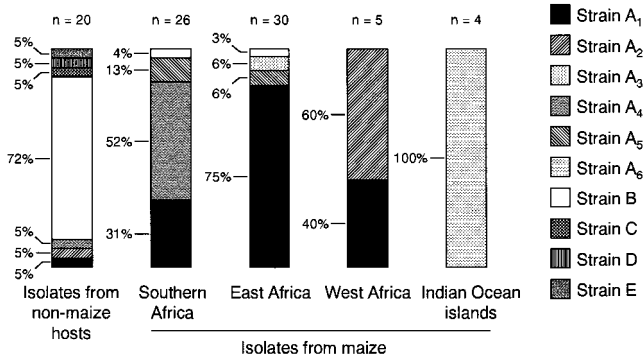
FIG. 2. The geographical and host distribution of 85 MSV isolates that have currently been typed in this and other MSV diversity studies (Briddon *et al.,* 1994; Rybicki *et al.,* 1998; Willment *et al.,* 2001). While 54% of the isolates used to construct this figure were typed exclusively by PCR–RFLP analysis, the rest were typed by partial and full-length genomic sequencing (a description of the subtype and strain classifications of all these isolates can be obtained from our Web site: ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/tablea.pdf). Numbers above bars (*n*) indicate the number of isolates used to construct the bars. Note that samples from nonmaize hosts were predominantly obtained from South Africa and percentages presented here are probably not representative of any individual nonmaize host species or of nonmaize hosts throughout the entire African continent.

quite rare and collectively comprised only 15% of isolates obtained from nonmaize hosts.

The overwhelming majority of MSVs obtained from maize were strain A isolates (Fig. 2). There appear to be substantial differences in the subtype composition of the MSV-A populations infecting maize in different parts of Africa (Fig. 2). While subtype A_1 isolates occur throughout Africa from Nigeria to South Africa, subtype A_2, A_3, and A_4 MSVs were found only in western, eastern, and

southern Africa, respectively (Fig. 2). Only subtype A_6 isolates have been detected in MSV-infected maize on the Indian Ocean islands of La Réunion and Mauritius (Briddon *et al.,* 1994) and no MSV-A_6 isolates have yet been detected on the African continent.

Pathogenicity of virus isolates in *Zea mays*

To determine the relationship between the genotypes of streak virus isolates and their pathogenicity in maize, we agroinoculated differentially MSV-resistant maize genotypes with 21 fully sequenced MSV, PanSV, DSV, SSRV, and SSV isolates. MSV strain A isolates produced the most severe symptoms in all the maize genotypes tested (Fig. 3). While all these isolates share greater than 95% nucleotide sequence identity, there are significant differences in their pathogenicity in maize. Whereas subtype A_1 (MSV-Ama, -Gat, -MatA, -MatB, and -Sag), A_2 (MSV-Ns), and A_5 (MSV-MakD and -MatC) isolates produced the severest symptoms in all the maize genotypes (C_{4-6} of 68.6–75.5), subtype A_4 isolates (MSV-Kom and -VM) produced the least severe symptoms (C_{4-6} of 54.0–54.4). Although subtypes A_3 (MSV-MtKA) and A_6 (MSV-R2) isolates generally produced intermediate symptoms in all the maize genotypes (C_{4-6} of 62.8–63.2), the subtype A_6 isolate produced symptoms as severe as the subtype A_1, A_2, and A_5 isolates in the MSV-resistant maize genotype PAN6099 (Fig. 3).

The MSV-B, -C, -D, and -E isolates were all substantially less severe (C_{4-6} of 32.3–5.5) than the MSV-A isolates in all the maize genotypes. The single maize-derived strain B isolate (MSV-Mom) was significantly less severe in maize than the maize-derived strain A isolates (Fig. 3). No strain B, C, D, or E isolates produced any
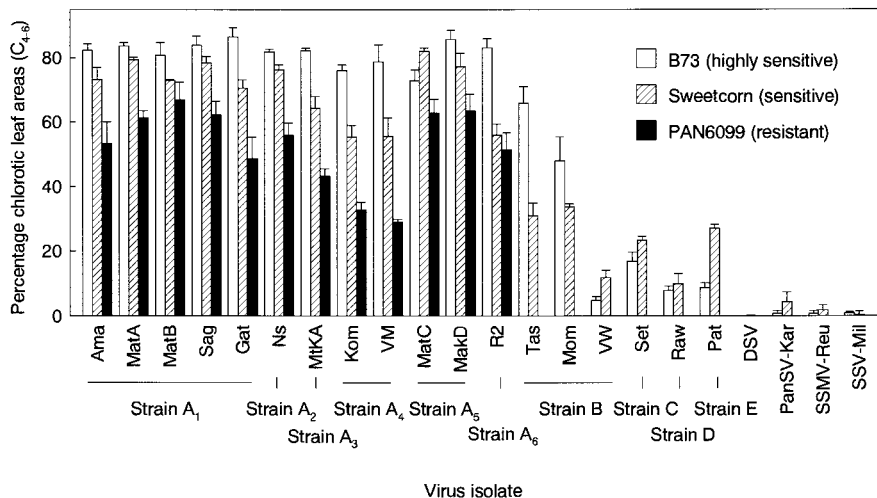

FIG. 3. Average percentages of chlorotic leaf areas occurring on leaves four through six (C_{4-6}) of different maize genotypes agroinfected with a range of MSV, PanSV, SSV, SSMV, and DSV isolates. The MSV isolates have been labeled according to their strain and subtype groupings. Maize genotypes were highly sensitive (B73, white bars), moderately sensitive (sweet corn, hatched bars), or moderately resistant (PAN6099, black bars) to MSV infection. No agroinfectious clones were available for the isolates presented in Fig. 1 that are not represented here (MSV-K, -SA, -Jam, PanSV-Ken, SSEV-Asw, -Ben, and SSV-N). Error bars represent 95% confidence intervals of the means (based on Student's *t* distribution).
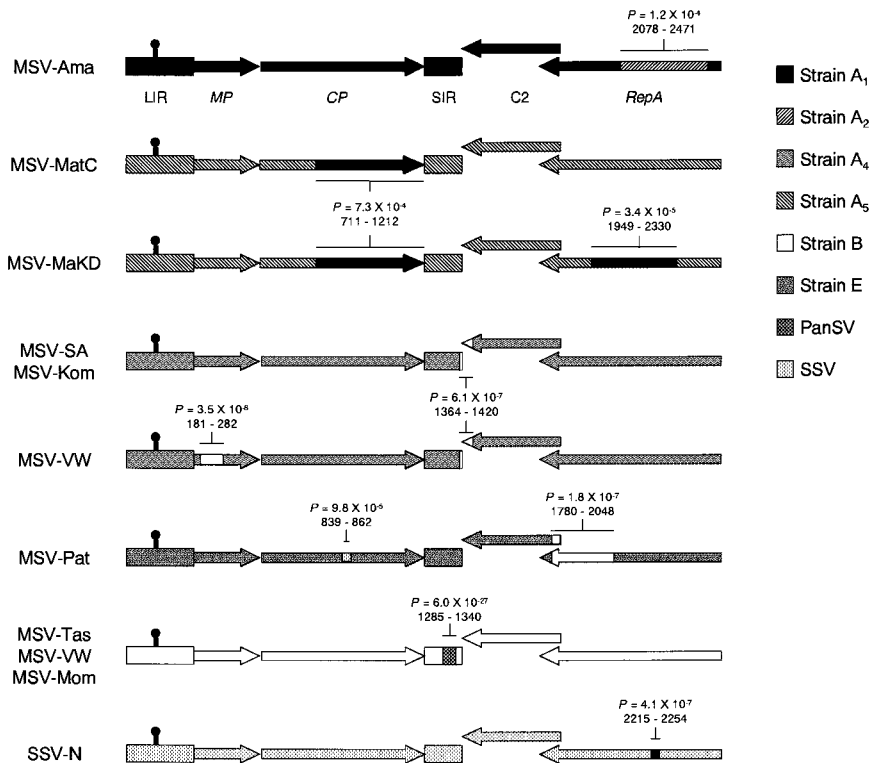
FIG. 4. A schematic representation of the recombinant regions in the African streak virus genomes analyzed during this study. Indicated are the major genomic features of all the recombinant virus isolates that were detected (labeled for MSV-Ama). LIR, long intergenic region; *MP*, movement protein gene; *CP*, coat protein gene; SIR, short intergenic region; *RepA* + C2, replication associated protein gene. The origin of different genomic regions is indicated by their shading. Note that while the recombinant region detected within *RepA* of SSV-N is presented here as being MSV-A_1-like, it could be more accurately described as being MSV-A-like. Potential recombination break points and the probability (P) that the indicated regions do not have a recombinant origin (Martin and Rybicki, 2000) are presented.

symptoms in the moderately resistant maize genotype PAN6099.

The PanSV, DSV, SSMV, and SSV isolates produced the mildest symptoms in maize (C_{4-6} of 0.06 to 1.69). None of these viruses produced any symptoms in PAN6099. Whereas the PanSV isolate was slightly more severe than the rest of these viruses (particularly in sweet corn), DSV was by far the mildest and produced no symptoms in B73.

Detection of recombination between virus isolates

Because recombination has potentially resulted in the recent emergence of a number of serious begomovirus diseases (Umaharan *et al.,* 1998; Zhou *et al.,* 1997, 1998), the prevalence of recombination among MSV genomes was investigated. Using a custom-written analysis software (RDP version 1.8; Martin and Rybicki, 2000), nine unique recombination events were detected among 28 full-length African streak virus sequences (Fig. 4). Apparently recombinant regions ranged in size from 23 to 501 nucleotides, with interspecies recombination apparently involving smaller regions (23–55 nucleotides) than intraspecies recombination (56–501 nucleotides).

Two groups of recombinant sequences were particu-

larly interesting. All the subtype A_4 sequences examined (MSV-SA, -Kom, and -VW) contain strain B-like sequences at the 3′ termini of their C2 ORFs (Fig. 4). This potentially recombinant region has apparently altered the position of the C2 termination codon so that subtype A_4 viruses would be predicted to express Rep proteins that, relative to viruses belonging to the other MSV-A subtypes, are missing five C-terminal amino acids. The MSV-B isolates MSV-Mom, -Tas, and -VW all contain 55 nucleotides of PanSV-like sequence in their short intergenic regions (SIRs). The recombinant portion of their SIRs is within the ~80-nucleotide region to which a primer-like oligonucleotide involved in virus replication has been found bound in encapsidated virions (Donson *et al.,* 1984).

DISCUSSION

Whereas a number of studies have begun to reveal the extent of MSV diversity (Briddon *et al.,* 1994; Hughes *et al.,* 1992; Rybicki *et al.,* 1998; Schnippenkoetter *et al.,* 2001; Willment *et al.,* 2001), virtually nothing is known of either the geographical and host distributions or the relative pathogenic potentials of the virus genotypes that have been detected. Our aim here was both to provide

an expanded view of MSV diversity and to explore the relationships between the known MSV genotypic groupings and their host specificities, geographical distributions, and virulence in maize.

For descriptive purposes, 22 unique full-length MSV genome sequences were classified into strain and subtype groupings. While these groupings were not intended to be definitive, they are supported directly by phylogenetic analysis of the sequences and indirectly by both their pathogenicity in differentially MSV resistant maize genotypes and their geographical distributions.

It is potentially controversial that the isolates classified in strains MSV-C, -D, and -E were grouped together in the same species as MSV-A and -B isolates, since they share less nucleotide sequence identity (between 78 and 83%) than is shared by many distinct geminivirus species in the genus *Begomovirus*. However, it has been proposed that a species be defined as any group of isolates capable of *trans*-replicating one another (Rybicki *et al.,* 2000). The *trans*-replication of MSV-A and -C isolates has been demonstrated (Schnippenkoetter, 1998; Willment, 1999). The fact that these strains are more dissimilar than are MSV-A and either MSV-D or -E implies that MSV-A isolates may also be able to *trans*-replicate and be *trans*-replicated by both MSV-D and -E isolates. Moreover, the MSV-C, -D, and -E isolates we examined are all more virulent in maize than the distinct mastreviruses PanSV, DSV, and SSV, meaning that there is a less clear-cut biological difference between them and the MSV-A isolates than there is between the "MSVs" and other mastreviruses.

MSV is known to infect over 80 wild and cultivated grass species (Büchen-Osmond, 1998; Damsteegt, 1983; Konate and Traore, 1992). This and other studies have indicated that severe streak disease in one of these species, *Z. mays,* is predominantly caused by MSV-A isolates (Briddon *et al.,* 1994; Willment *et al.,* 2001). Collectively, all major MSV diversity studies that have typed MSV isolates by either partial sequencing or RFLP analysis have examined only 23 isolates from eight nonmaize host species, almost all of which have originated from South Africa. Given such a small sample size, it is interesting that five unique MSV strains have been detected in these hosts. It is not inconceivable that the overwhelming predominance of MSV-A isolates in maize is the consequence of a common MSV host adaptation strategy and that there may be nearly as many host-adapted MSV strains as there are host species.

While small, diverse, and potentially unrepresentative samples and host adaptation may explain why only single MSV-C, -D, and -E isolates have been detected, the relatively frequent occurrence of MSV-A and -B isolates within a wide range of grass species demands an explanation. Whereas MSV-A isolates have been detected in three nonmaize grass species (*Coix lachryma jobi,* a *Pennisetum* sp., and a *Digitaria* sp.; Briddon *et al.,* 1994),

MSV-B isolates have been detected in at least seven species (maize, wheat, rye, *Digitaria* spp., a *Paspalum* sp., an *Eleusine* sp., a *Setaria* sp., and *Urochloa* spp.; Willment *et al.,* 2001). MSV-A and -B are also the only MSV strains that have currently been detected in exotic cultivated grass species such as maize and wheat, which may indicate that they have a fundamentally different survival strategy from that of the other MSV strains. The ability of MSV-A and MSV-B isolates to effectively use a wide range of grass species as hosts was possibly an important factor enabling them to successfully infect and produce severe symptoms in different cultivated species that have been introduced into Africa from other parts of the world.

While too few MSV-B, -C, -D, and -E isolates were analyzed to determine the geographical distribution of subtypes within these strains, enough MSV-A samples were characterized to obtain a low-resolution picture of subtype distributions in three different parts of Africa. The six MSV-A subtypes identified each displayed varying degrees of geographical localization. Two of the subtypes, MSV-A_1 and MSV-A_5, are widely distributed, with both being found in eastern and southern Africa and MSV-A_1 also in west Africa. Currently MSV-A_2, MSV-A_3, and MSV-A_4 isolates have only been found in west Africa, east Africa, and southern Africa, respectively. MSV-A_6 isolates have only been found on the Indian Ocean islands of Mauritius and La Réunion.

We detected potentially important differences between the virulence of different MSV-A subtypes in maize. The degree of symptom severity produced by the MSV-A_1, -A_2, and -A_5 isolates was greater than that produced by the MSV-A_3 and -A_4 isolates we examined. It is possible that the wider distribution of MSV-A_1 and MSV-A_5 relative to that of other MSV-A subtypes is related to their greater virulence in maize.

The virulence of all the non-MSV-A isolates tested appears to be roughly correlated with the degree of nucleotide sequence identity they share with MSV-A isolates. Therefore, while MSV-B, -C, -D, and -E isolates were less virulent than MSV-A isolates, all MSV isolates were more virulent than other African streak viruses. Virtually every MSV isolate detected to date has been obtained from either cultivated or wild annual grass species. With the exception of SSV-Mil (Briddon *et al.,* 1996), every non-MSV African streak virus currently characterized has been obtained from perennial grass species. It is possible that adaptation of MSV isolates to infecting annual grass species—which perhaps requires the virus to produce more striking symptoms more quickly, to guarantee leafhopper transmission—has predisposed them to greater virulence in maize than the other African streak virus species.

It is unlikely that non-MSV-A isolates could pose any direct threat to maize production. MSV-A isolates tested were all significantly more severe in maize than all other

virus isolates we examined and only MSV-A isolates were shown to be capable of producing any symptoms in the moderately MSV-resistant maize genotype PAN6099. Of the MSV-A subtypes, A_1 and A_5 potentially pose the most significant threat to maize production because of their wide distribution throughout Africa and their extreme virulence relative to most of the other MSV-A isolates tested.

Because widespread recombination has been detected among geminiviruses in the genus *Begomovirus* (Padidam *et al.,* 1999) and has potentially been responsible for the emergence of highly pathogenic virus genotypes (Umaharan *et al.,* 1998; Zhou *et al.,* 1997, 1998), we examined all available African streak virus sequences for evidence of recombination. An extensive study of recombination among geminiviruses indicated that recombination among mastreviruses occurs at a much lower frequency than that which occurs among begomoviruses (Padidam *et al.,* 1999). Our results indicate that this observation may not be accurate: enough MSV sequences were examined here to detect a substantial degree of recombination both within this species and with other African streak viruses. There are, however, still too few non-MSV African streak virus sequences available to properly detect recombination between them and any species other than MSV.

It is unknown whether any of the recombinant viruses we identified have modified host ranges or enhanced virulence under certain conditions. None of the recombinants was detectably more pathogenic in maize than closely related nonrecombinant viruses. In three cases, nearly identical recombinant regions were noted in two or more isolates obtained from different locations and in different years, indicating that certain recombinants detected in this study are potentially major circulating MSV forms. Whether recombination has provided any of these viruses with subtle selective advantages over their nonrecombinant counterparts is, however, unknown and may be difficult to test experimentally if, for example, these advantages are only apparent in certain host species or cell types.

Through an analysis of MSV genotypic diversity, distribution, and virulence, we have identified MSV-A subtypes that could potentially pose the greatest threat to maize production in Africa. It is hoped that our preliminary low-resolution description of MSV-A genotype distributions will aid future epidemiological investigations of changes in MSV-A population compositions over time. Whereas it was determined that only MSV-A isolates were likely to have any direct impact on maize production, the identification of inter-MSV strain recombination indicates that strains other than MSV-A may have an indirect impact on the epidemiology of maize streak disease. It has been widely speculated that recombination may be responsible for the emergence of certain begomoviral diseases (Padidam *et al.,* 1999; Umaharan

et al., 1998; Zhou *et al.,* 1997, 1998) and it is quite conceivable that recombination might also yield MSV variants with enhanced pathogenic properties.

MATERIALS AND METHODS

Maize genotypes and virus isolates

Seed for the maize genotypes B73 and PAN6099 was provided by J. B. J. van Rensburg (Summer Grains Centre, Potchefstroom, South Africa) and D. Nowell (PANNAR Ltd., Greytown, South Africa), respectively. Sweet corn (cv. Jubilee) seed was obtained from Starke Ayres Ltd. (Cape Town, South Africa). A list of the 85 MSV isolates involved in this study can be obtained from our Web site (ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/tablea.pdf). The origin of cloned full-length and completely sequenced virus genomes used to produce agroinfectious constructs are presented in Table 1.

Detection of MSV diversity

DNA from the leaves of 46 MSV infected maize (37 samples collected between 1994 and 2000 from five locations in southern Africa, 24 locations in east Africa, and 1 location in west Africa), *Urochloa* spp., an *Eleusine* sp., and *Digitaria* spp. plants (collectively 9 samples collected between 1998 and 2000 from 6 locations in southern Africa and 1 in east Africa) was isolated by the method of Palmer *et al.* (1998) and analyzed using the PCR–RFLP MSV typing technique of Willment *et al.* (2001). For each of the MSV isolates examined, this analysis involved (1) the PCR amplification (using the degenerate primer set 5′-CAAAKDTCAGCTCCTCCG-3′ and 5′-TTGGVCCGMVGATGTASAG-3′) of a ~1300-nucleotide sequence spanning *RepA,* the LIR, and the 5′ half of *MP,* (2) digestion of the PCR product with the restriction enzymes *Rsa*I, *Hpa*II, *Hae*III, *Sau*IIIa, *Bam*HI, *Cfo*I, and *Hin*dIII (Boehringer Mannheim, Mannheim, Germany), and (3) matching restriction fragment lengths to those of a reference set of fully or partially sequenced isolates (Willment *et al.,* 2001).

Cloning of full-length MSV genomes

MSV replicative form (RF) DNAs were isolated as described by Palmer *et al.* (1998). RF DNAs were linearised using either *Bam*HI (MSV-MatA, MSV-MatB, MSV-MatC, MSV-MtKA, MSV-Sag, MSV-Ama, MSV-Gat, MSV-MakD, MSV-Jam, and MSV-Pat) or *Sal*I (MSV-Mom and MSV-Raw) and cloned using standard techniques (Sambrook *et al.,* 1989) into similarly linearized pBluescript(SK$^+$) (pSK$^+$; Stratagene, La Jolla, CA).

Sequencing and analysis of full-length MSV genomes

Sequencing of the 12 cloned MSV isolates was carried out using an ALF Express automated sequencer (Pharmacia Corp., Peapack, NJ). The full nucleotide se-

TABLE 1

Agroinfectious *Mastrevirus* Isolates Involved in This Study

Virus isolate	Source of isolate	Year of isolation	Region and country of origin	Reference
MSV-Ama	*Zea mays*	1998	Amagoro, Kenya	This study
MSV-Gat	*Zea mays*	1998	Gathuke-ini, Kenya	This study
MSV-Kom	*Zea mays*	1989	Komatipoort, South Africa	Schnippenkoetter *et al.* (2001)
MSV-MakD	*Zea mays*	1998	Makatini, South Africa	This study
MSV-MatA	*Zea mays*	1994	Matabeleland, Zimbabwe	This study
MSV-MatB	*Zea mays*	1996	Matabeleland, Zimbabwe	This study
MSV-MatC	*Zea mays*	1998	Matabeleland, Zimbabwe	This study
MSV-Mom	*Zea mays*	1998	Mombasa, Kenya	This study
MSV-MtKA	*Zea mays*	1997	Mt. Kenya, Kenya	This study
MSV-Ns	*Zea mays*	1984	Nigeria	Mullineaux *et al.* (1984)
MSV-R2	*Zea mays*	1997	La Réunion	Isnard *et al.* (1998)
MSV-Sag	*Zea mays*	1998	Sagana, Kenya	This study
MSV-VM	*Zea mays*	1993	Vaalharts, South Africa	Willment (1999)
MSV-Jam	*Digitaria* sp.	1999	Jamaica, Kenya	This study
MSV-Pat	*Digitaria* sp.	1999	Durban, South Africa	This study
MSV-Raw	*Urochloa* sp.	1998	Rawsonville, South Africa	This study
MSV-Set	*Setaria* sp.	1988	Durban, South Africa	Schnippenkoetter *et al.* (2001)
MSV-Tas	*Triticum aestivum*	1991	Elseberg, South Africa	Willment (1999)
MSV-VW	*Triticum aestivum*	1993	Vaalharts, South Africa	Willment (1999)
DSV	*Digitaria sanguinalis*	1987	Vanuatu	Donson *et al.* (1987)
PanSV-Kar	*Panicum maximum*	1989	Karina, South Africa	Schnippenkoetter *et al.* (2001)
SSV-Mil	*Pennisetum* sp.	1994	Kenya	Briddon *et al.* (1996)
SSRV-Reu	*Saccharum officinarum*	1998	La Réunion	Bigarré *et al.* (1999)

quences of these isolates were determined in both orientations. Sequence assembly, alignment, and phylogenetic analyses were done using DNAMAN (version 4.0; Lynnon BioSoft, Quebec) and PHYLIP (version 3.57c; J. Felsenstein, University of Washington). Open reading frames and intergenic regions within newly sequenced genomes were inferred by analogy with those of other MSV isolates. Nucleotide sequence elements and amino acid motifs with either known or suspected functions within predicted gene products were also inferred by analogy with other mastreviruses (reviewed by Palmer and Rybicki, 1998; Gutierrez, 1999). Annotated alignments of all the nucleotide and amino acid sequences determined in this study can be found on our Web site (ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/figc. pdf).

Generation of agroinfectious constructs

Agroinfectious clones were constructed according to the method of Palmer (1997). *Xba*I–*Eco*RI fragments containing full tandem repeats of MSV-MatA, MSV-MatB, MSV-MatC, MSV-MtK, MSV-Sag, MSV-Ama, MSV-Gat, and MSV-MakD genomes were cloned from pSK$^+$ into the *Xba*I and *Eco*RI sites of pBI121 (CLONTECH, CA). *Xba*I–*Xho*I fragments containing full tandem repeats of MSV-Pat, MSV-Mom, and MSV-Raw genomes were cloned from pSK$^+$ into the *Xba*I–*Sal*I sites of pBIN19 (CLONTECH). Agroinfectious clones were transformed into *Agrobacterium tumefaciens* C58C1 (pMP90) (Koncz and Schell, 1986) using the method of An *et al.* (1988).

Agroinfectious virus constructs were also obtained from M. I. Boulton (MSV-Ns, DSV; John Innes Centre, Norwich, UK), M. Peterschmitt (MSV-R2X2, SSRV-Reu; CIRAD, Montpellier, France), R. Briddon (SSV-Mil; John Innes Centre), W. Schnippenkoetter (MSV-Kom, MSV-Set, PanSV-Kar; University of Cape Town, Cape Town, South Africa), and J. Willment (MSV-VM, MSV-VW, MSV-Tas; University of Cape Town) (Table 1).

Agroinoculation and evaluation of isolate pathogenicity

Agroinoculation was performed according to Grimsley *et al.* (1987) with modifications described previously (Martin and Rybicki, 1998). Agroinfectious constructs of all the MSV, DSV, SSV, and PanSV isolates were used to infect three maize genotypes—B73 (highly susceptible), sweet corn (moderately susceptible), and PAN6099 (moderately resistant). Agroinfection experiments involving each maize genotype–MSV isolate combination were repeated at least three times. Agroinfection experiments involving specific constructs were carried out on groups of 14 plants.

Percentages of leaf areas covered by chlorotic lesions in symptomatic plants were estimated for leaves four through six using a microcomputer-based image analysis technique described previously (Martin and Rybicki, 1998; Martin *et al.,* 1999). Percentages of chlorotic areas of leaves four, five, and six were assessed 22, 29, and 35 days after agroinoculation, respectively. For each virus isolate–plant genotype combination, the mean percent-

ages of chlorotic areas of the fourth to sixth leaves (designated C_{4-6}) were used as representative measures of chlorosis. For each MSV isolate, a single value designated C_{4-6}, representing the mean of the C_{4-6} values calculated for B73, sweet corn, and PAN6099, was used as a generalized measure of pathogenicity.

Detection of recombinant genomes

Detection of potentially recombinant sequences, identification of likely parent sequences, and localization of possible recombination break points was carried out using RDP (Recombination Detection Program; Martin and Rybicki, 2000). The software has been developed in this laboratory and is available free of charge from our Web site (http://www.uct.ac.za/depts/microbiology/microdescription.htm). The RDP settings used were multiple comparison correction off, internal reference selection, and a window size of 10. An alignment containing 28 full-length African streak virus sequences was analyzed for evidence of intersubtype, interstrain, and interspecies recombination in two sets, each using a different highest acceptable P-value setting. The first set contained 15 MSV-A sequences (MSV-Ns, MSV-MatA, MSV-MatB, MSV-MatC, MSV-MakD, MSV-Ama, MSV-Sag, MSV-Gat, MSV-MtKA, MSV-Ken, MSV-Kom, MSV-SA, MSV-VM, MSV-Reu, and MSV-R2) and was analyzed with a highest acceptable P value of 1×10^{-3}. The second set contained 22 MSV (the 15 MSV-A sequences, MSV-Tas, MSV-Mom, MSV-VW, MSV-Jam, MSV-Pat, MSV-Set, MSV-Raw), 2 PanSV (PanSV-Ken and PanSV-Kar), and 4 SSV sequences (SSEV-Asw, SSEV-Ben, SSMV-Reu, and SSV-N; GenBank accession numbers for sequences not determined in this study are available from our Web site: ftp://ftp.uct.ac.za/pub/data/geminivirus/diversity/tableb.pdf) and was analyzed with a highest acceptable P value of 1×10^{-4}.

Neighbor joining trees (Saitou and Nei, 1987) were constructed using regions within the aligned genomes that RDP identified as having a potentially recombinant origin and in all cases bootstrap analysis was used to confirm the program's findings. Briefly, this involved realignment of sequences within all regions of the original alignment that corresponded with sets of potential recombination break points and construction of a series of neighbor joining trees, each with 1000 bootstrap iterations. Acceptable confirmation that a sequence was of recombinant origin was "movement" of the sequence with over 70% bootstrap support from its position in a tree constructed with full-length sequences to a divergent branch of a tree constructed using only a potentially recombinant region within the alignment.

ACKNOWLEDGMENTS

REFERENCES

An, G., Ebert, P. R., Mitra, A., and Ha, S. B. (1988). Binary vectors. *In* "Plant Molecular Biology Manual" (S. B. Gelvin and R. A. Schilperoort, Eds.), pp. 87 and 88. Kluwer, Dordrecht, The Netherlands.

Bigarré, L., Salah, M., Granier, M., Frutos, R., Thouvenel, J.-C., and Peterschmitt, M. (1999). Nucleotide sequence evidence for three distinct sugarcane streak mastreviruses. *Arch. Virol.* **144,** 2331–2344.

Bosque-Perez, N. A. (2000). Eight decades of *Maize streak virus* research. *Virus Res.* **71,** 107–121.

Briddon, R. W., Lunness, P., Chamberlin, L. C., Pinner, M. S., Brundish, H., and Markham, P. G. (1992). The nucleotide sequence of an infectious insect-transmissible clone of the geminivirus Panicum streak virus. *J. Gen. Virol.* **73,** 1041–1047.

Briddon, R. W., Lunness, P., Chamberlin, L. C., and Markham, P. G. (1994). Analysis of the genetic variability of maize streak virus. *Virus Genes* **9,** 93–100.

Briddon, R. W., Lunness, P., Bedford, I. D., Chamberlin, L. C., Mesfin, T., and Markham, P. G. (1996). A streak disease of pearl millet caused by a leafhopper-transmitted geminivirus. *Eur. J. Plant Pathol.* **102,** 397–400.

Büchen-Osmond, C. (1998). ICTVdB. http://www.ncbi.nlm.gov/ICTVdB/ICTV/dB/ 29010006.htm, *Maize streak virus*.

Castellano, M. M., Sanz-Burgos, A. P., and Gutierrez, C. (1999). Initiation of DNA replication of a eukaryotic rolling-circle replicon: Identification of multiple DNA-protein complexes at the geminivirus origin. *J. Mol. Biol.* **290,** 639–652.

Damsteegt, V. D. (1983). Maize Streak Virus: I. Host range and vulnerability of maize germplasm. *Plant Dis.* **67,** 734–737.

Donson, J., Morris-Krsinich, B. A. M., Mullineaux, P. M., Boulton, M. I., and Davies, J. W. (1984). A putative primer for second-strand DNA synthesis of maize streak virus is virion-associated. *EMBO J.* **3,** 3069–3073.

Donson, J., Accotto, G. P., Boulton, M. I., Mullineaux, P. M., and Davies, J. W. (1987). The nucleotide sequence of a geminivirus from *Digitaria sanguinalis. Virology* **161,** 160–169.

Gutierrez, C. (1999). Geminivirus DNA replication. *Cell. Mol. Life Sci.* **56,** 311–329.

Hughes, F. L., Rybicki, E. P., and von Wechmar, M. B. (1992). Genome typing of Southern African subgroup-1 geminiviruses. *J. Gen. Virol.* **73,** 1031–1040.

Isnard, M., Granter, M., Frutos, R., Reynaud, B., and Peterschmitt, M. (1998). Quasi-species nature of the maize streak virus isolates obtained from a population used to assess maize cultivar response to infection. *J. Gen. Virol.* **79,** 3091–3099.

Konate, G., and Traore, O. (1992). Reservoir hosts of maize streak virus (MSV) in the Sudan-Sahel zone: Identification and spatio-temporal distribution. *Phytoprotection* **73,** 111–117.

Koncz, C., and Schell, J. (1986). The promoter of T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204,** 383–396.

Kotlizky, G., Boulton, M. I., Pitaksutheepong, C., Davies, J. W., and Epel, B. L. (2000). Intracellular and intercellular movement of maize streak geminivirus V1 and V2 proteins transiently expressed as green fluorescent protein fusions. *Virology* **274,** 32–38, doi:10.1006/viro.2000.0415.

Liu, H., Boulton, M. I, Oparka, K. J., and Davies, J. W. (2001). Interaction of the movement and coat proteins of Maize streak virus: Implications for the transport of viral DNA. *J. Gen. Virol.* **82,** 35–44.

Liu, L., Pinner, M. S., Davies, J. W., and Stanley, J. (1999). Adaptation of the geminivirus bean yellow dwarf virus to dicotyledonous hosts

involves both virion-sense and complementary-sense genes. *J. Gen. Virol.* **80,** 501–506.

Martin, D. P., and Rybicki, E. P. (1998). Microcomputer based quantification of maize streak virus symptoms in *Zea mays. Phytopathology* **88,** 422–427.

Martin, D. P., Willment, J. A., and Rybicki, E. P. (1999). Evaluation of maize streak virus pathogenicity in differentially resistant *Zea mays* genotypes. *Phytopathology* **89,** 695–700.

Martin, D. P., and Rybicki, E. P. (2000). RDP: Detection of recombination amongst aligned sequences. *Bioinformatics* **16,** 562–563.

Missich, R., Ramirez-Parra, E., and Gutierrez, C. (2000). Relationship of oligomerisation to DNA binding of *Wheat dwarf virus* RepA and Rep proteins. *Virology* **273,** 178–188, doi:10.1006/viro.2000.0412.

Morra, M. R., and Petty, I. T. (2000). Tissue specificity of geminivirus infection is genetically determined. *Plant Cell* **12,** 2259–2270.

Mullineaux, P. M., Donson, J., Morris-Krsinich, B. A., Boulton, M. I., and Davies, J. W. (1984). The nucleotide sequence of maize streak virus DNA. *EMBO J.* **3,** 3063–3068.

Padidam, M., Sawyer, S., and Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265,** 218–225, doi:10.1006/viro.1999.0056.

Palmer, K. E. (1997). "Investigations into the Use of *Maize streak virus* as a Gene Vector." Ph.D. thesis, University of Cape Town.

Palmer, K. E., and Rybicki, E. P. (1998). The molecular biology of mastreviruses. *Adv. Virus Res.* **50,** 183–234.

Palmer, K. E., Schnippenkoetter, W. H., and Rybicki, E. P. (1998). Geminivirus isolation and DNA extraction. *In* "Plant Virology Protocols: From Virus Isolation to Transgenic Resistance" (G. Foster and S. Taylor, Eds.). Humana Press, Totawa.

Peterschmitt, M., Granier, M., Frutos, R., and Reynaud, B. (1996). Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Réunion Island. *Arch. Virol.* **141,** 1637–1650.

Petty, I. T., Carter, S. C., Morra, M. R., Jeffrey, J. L., and Olivey, H. E. (2000). Bipartite geminivirus host adaptation determined cooperatively by coding and noncoding sequences of the genome. *Virology* **277,** 429–438, doi:10.1006/viro.2000.0620.

Rybicki, E. P., Dennis, S., Napier, G., and Hughes, F. L. (1998). Novel grass and wheat strains of maize streak virus detected by DNA amplification and sequencing. *Mol. Plant Pathol. On-Line* http://194.247.68.33/mppol/1998/0929rybicki.

Rybicki, E. P., Briddon, R. W., Brown, J. E., Fauquet, C. M., Maxwell, D. P., Harrison, B. D., Markham, P. G., and Stanley, J. (2000). Virus taxonomy. *In* "Seventh Report of the International Committee on Taxonomy of Viruses" (C. M. Fauquet, D. H. L. Bishop, E. B. Carsten, M. K. Estes, S. M. Lemon, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner, Eds.), pp. 285–297. Academic Press, San Diego.

Saitou, N., and Nei., M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schaffer, R. L., Miller, C. G., and Petty, I. T. (1995). Virus and host-specific adaptations in the BL1 and BR1 genes of bipartite geminiviruses. *Virology* **214,** 330–338.

Schnippenkoetter, W. H., Martin D., Hughes, F., Fyvie, M., Willment, J., James, D., von Wechmar, B., and Rybicki, E. P. (2001). The relative infectivities and genomic characterisation of three mastreviruses. *Arch. Virol.* **146,** 1075–1088.

Umaharan, P., Padidam, M., and Phelps, R. H. (1998). Distribution and diversity of geminiviruses in Trinidad and Tobago. *Phytopathology* **88,** 1262–1268.

Willment, J. A. (1999). "Investigation of the Molecular Determinants of *Maize streak virus* Replication." Ph.D. thesis, University of Cape Town.

Willment, J. A., Martin, D. P., and Rybicki, E. P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLP's and partial sequence data. *J. Virol. Methods* **93,** 75–87.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J., and Harrison, B. D. (1997). Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* **78,** 2101–2111.

Zhou, X., Liu, Y., Robinson, D. J., and Harrison, B. D. (1998). Four DNA-A variants among Pakistani isolates of cotton leaf curl virus and their affinities to DNA-A of geminivirus isolates from okra. *J. Gen. Virol.* **79,** 915–923.