Quasispecies nature of three maize streak virus isolates obtained through different modes of selection from a population used to assess response to infection of maize cultivars

Muriel Isnard,1 Martine Granier,1 Roger Frutos,1 Bernard Reynaud2 and Michel Peterschmitt1

1 Programme Protection des Cultures, Laboratoire IGEPAM, CIRAD-AMIS, Bât. II, BP 5035, 34032 Montpellier Cedex 1, France
2 CIRAD-AMIS, 7 Chemin de l’IRAT, 97410 Saint Pierre, La Réunion, France

Three maize streak virus (MSV) isolates were derived from an MSV population used to assess the response to infection of maize cultivars. Isolate SP1 was obtained from this population through short acquisition and inoculation periods (1 and 5 min, respectively), using a single Cicadulina mbila vector. Isolate SP2 was derived from SP1 after transmission to a wild perennial host (Coix lacryma-jobi), on which it was maintained for about 4 years without insect transmission. Isolate N2A, the most pathogenic isolate, was obtained from the initial population after serial passages on almost completely resistant inbred maize lines. The complexity of each isolate was analysed by RFLP analysis and sequencing based on 120 SP1 clones, 36 SP2 clones and 40 N2A clones. All three isolates were composed of different but related clones, consistent with a quasispecies structure. The mutations were distributed throughout the genome. Mutation frequencies, based on all available sequences, were $3.8 \times 10^{-4}$ for SP1, $1.05 \times 10^{-4}$ for SP2 and $6.9 \times 10^{-4}$ for N2A. As expected from the bottleneck selection step, the intra-isolate variability of SP1 was relatively low. Comparison between SP1 and SP2 showed that SP1 heterogeneity increased during maintenance on the wild host. Furthermore, the consensus sequences of SP1 and SP2 differed by two non-synonymous substitutions in the complementary sense gene repA. N2A had a relatively low degree of heterogeneity, but was composed of several sub-populations. The results reflect the influence of the mode of selection of MSV isolates on their quasispecies organization, i.e. distribution of variants, and master sequence.

Introduction

Maize streak virus (MSV) is a member of the family Geminiviridae (Briddon & Markham, 1995), which is composed of plant viruses with geminate quasi-icosahedral particles and circular single-stranded DNA genomes (for review see Lazarowitz, 1992). It is the type member of the genus Mastrevirus, characterized by leafhopper transmission, mainly monocotyledonous hosts and a monopartite genome. MSV is widely distributed across sub-Saharan Africa as well as on Indian Ocean islands (Thottappilly et al., 1993; Peterschmitt et al., 1991). Five MSV clones have been sequenced fully, isolated in Nigeria (Mullineaux et al., 1984; Boulton et al., 1991), Kenya (Howell, 1984), South Africa (Lazarowitz, 1987) and La Réunion (Peterschmitt et al., 1996).

Molecular studies show low variability among MSV isolates from maize. Although some variations among African isolates were detected by RFLP analysis (Hughes et al., 1992), these isolates were classified with island isolates (La Réunion and Mauritius) into the same serotype on the basis of analysis with monoclonal and polyclonal antibodies (Pinner et al., 1988; Dekker et al., 1988; Peterschmitt et al., 1991). This is consistent with the high similarity, ranging from 98 to 100%, of the coat protein (CP) amino acid sequences of 12 MSV isolates from different geographical locations (Briddon et al., 1994). How-
ever, on the basis of partial and total genome sequences, it was shown that the Mauritian and La Réunion isolates are distinct from African isolates (Briddon et al., 1994; Peterschmitt et al., 1996); about 4–1% variation for total sequences was observed. The genome variability between clones from different African isolates (2.0–2.2%) is similar to that of European wheat dwarf virus (WDV) clones (1.3–1.4%) (Bendahmane et al., 1995), but much lower than that found within clones of another monopartite geminivirus, beet curly top virus (BCTV) (17.2–21%; Stenger & Ostrow, 1996). From an agronomic viewpoint, this apparent homogeneity of MSV suggests that resistant cultivars of maize should be effective in most African countries. However, it is not yet known how MSV isolates will evolve on resistant cultivars and whether resistance-breaking variants will appear. This risk cannot be ruled out, as it is known from other studies that viruses are able to adjust to new conditions by virtue of the quasispecies nature of their populations (for review, see Domingo et al., 1996). Several observations suggest that the MSV genome has a quasispecies nature; mild MSV isolates have been obtained from severe isolates through several vector feeding conditions (Pinner et al., 1988; Peterschmitt et al., 1991; Boulton et al., 1991), and nucleotide variations were detected by sequencing several clones belonging to the same isolate (Lazarowitz, 1987; Mullineaux et al., 1984; Howell, 1984). Furthermore, single nucleotide substitutions can change MSV severity (Boulton et al., 1991), indicating that spontaneous point mutations can have a significant effect on virus phenotype.

In this work we wanted first to confirm the quasispecies nature of MSV using three MSV isolates, SP1, SP2 and N2A, obtained through different modes of selection from a population used to assess the response to infection of maize cultivars (Rodier et al., 1995). The three isolates were shown to be of a quasispecies nature by RFLP analysis and sequencing. Secondly, since these three isolates were subjected to different environmental conditions (bottleneck, host change, plant resistance) for their selection, the influence of these conditions on the MSV quasispecies could be analysed by comparing the composition and the master sequence of the three populations.

Methods

- Virus isolates. MSV isolates SP1, SP2 and N2A were derived from an isolate used to screen maize cultivars for resistance to maize streak disease near Saint Pierre (La Réunion). This isolate is the progeny of a population obtained 10 years ago by pooling several isolates collected in different locations on the island. The isolate was maintained on partially resistant maize plants.

Isolate SP1 was obtained in 1991 through a single Cicadulina mbila vector transmission, with a 1 min acquisition on a maize plant infected by the screening isolate, followed by a 5 min inoculation to a susceptible healthy maize plant (Peterschmitt et al., 1996).

Two months after single-insect transmission, the SP1 isolate was transmitted by many vectors onto a perennial host, Coix lacryma-jobi, on which it was maintained for 4 years without any leafhopper transmission. Each year, young sprouts collected from the old plant were cultivated in a new pot. Before analysis, this isolate, SP2, was transmitted by many vectors onto susceptible maize plants of a temperate hybrid (Sabrina, Pioneer) to provide a source of plant material for DNA extraction.

Isolate N2A was obtained in 1993 from the screening isolate; the potential effect of plant resistance on the isolate was increased by passing it several times on highly, but not completely, resistant inbred lines, B422 and A211 (Rodier et al., 1995). The passages were as follows. Passage 1: inoculation of B422 plants with the screening isolate. One B422 plant rated 4 [on a scale from 1 to 5 used for assessing MSV symptoms on infected plants, according to Rodier et al. (1995)] was used as the source plant for inoculating the cultivar Sabrina, in which the isolate was maintained for about 5 months. Passage 2: infected Sabrina plants were used as source plants for inoculating healthy A211 plants. Passage 3: three infected A211 plants were used as source plants to inoculate 29 A211 plants; only one plant was infected (3.4% incidence). Passage 4: this plant was used as the source to inoculate 18 A211 plants, of which three became infected (16.7%). Passage 5: the three infected plants were used as the source to inoculate 80 A211 plants; 100% infection was obtained. For each passage, insects were kept for 2 or 3 days on infected leaves, and then transferred to healthy plants. Since many source plants exhibited mild symptoms, consisting of a few isolated chlorotic streaks of limited surface-area, the efficiency of transmission was increased by keeping vectors on these streaks in small cages clipped onto infected leaves. Isolate N2A was maintained on A211 plants for 10 months before use. To increase viral DNA concentration in infected plants, isolate N2A was transmitted by many insects to susceptible maize plants (Sabrina), from which DNA was extracted.

- DNA extraction, cloning and RFLP analysis. Replicative forms of N2A, SP1 and SP2 were each extracted from 3 g frozen leaves of MSV-infected maize plants. Extraction was carried out according to Hoisington et al. (1992), except the final step. After recovery of plant genomic DNA with a glass hook, the supernatant containing viral DNA was pelleted by centrifugation at 12000 r.p.m. for 30 min. The pellet, containing viral and sheared plant genomic DNA, was washed with 70% ethanol and resuspended in 500 µl sterile water. Cloning of viral DNA into pBC-KS (Stratagene) was done according to Peterschmitt et al. (1996).

RFLP analysis was performed on 40 clones of N2A, 120 of SP1 and 36 of SP2 using the following enzymes: AccI, AflII, BglII, EcoRV, HincII, HindIII, HpaI, SmaI, SacI and XhoI. These enzymes were chosen because of the distribution of their recognition sequences throughout the genomes of the five fully sequenced MSV clones.

- Sequence analysis. All nucleotide positions are given according to the sequence of MSV-R (Peterschmitt et al., 1996), called R1 in this study. A sample of clones of each isolate (11 from SP1, 10 from SP2 and 22 from N2A) was sequenced in a 430 bp region (positions 2255–2685) containing the large intergenic region (LIR), one of the most variable regions of the MSV genome (Peterschmitt et al., 1996), and the beginning of repA. Except for one SP1 clone, deleted in the LIR (see results), the same sample of SP1 and SP2 clones was further analysed to give sequences within CP (550–1020) and repA (1600–2255). Fifteen of the N2A clones were further analysed to provide sequences within the MP gene (1–100). For each isolate, at least one clone from each RFLP group was sequenced fully. Partially sequenced clones were designated with a number, whereas fully sequenced clones were designated by R followed by a number. Consensus sequences were derived for each isolate by alignment of partial and total sequences. Mutations were described by comparison to the consensus sequences.

Sequencing of cloned MSV DNA was carried out with the Taq dye terminator cycle sequencing kit (Applied Biosystems) and an Applied Biosystems 373A sequencer. Template DNA was prepared with Qiagen.
columns. Sequencing primers were either plasmid-specific primers (T3, T7) or internal MSV-specific primers. Sequences were stored, assembled and analysed with the Sequence Navigator and Auto Assembler software (Perkin-Elmer) and various applications from the Lasergene software (DNASTAR).

### Agroinfection

The infectivity of all fully sequenced MSV clones was tested, except for three SP2 clones (R11, R12 and R13) and two N2A clones (R5 and R9). Cloned MSV DNA was released from the BamHI cloning site of pBC-KS and purified from agarose gels with a Gene Clean kit (BIO 101). The MSV DNA was self-ligated for 10 min before adding dephosphorylated BamHI-cut pBC-KS to the ligation. Tandem dimers of MSV DNA were detected by restriction analysis. Selected recombinant plasmids were linearized by XhoI and end-filled with the Klenow fragment of DNA polymerase I. MSV tandem dimers were released with SalI, purified as above and ligated in a molar excess into SmaI–SalI-cut pBluescript (Bevan, 1984).

These constructs were introduced into Agrobacterium tumefaciens strain C58 C1, and the bacteria were resuspended in water and inoculated onto maize plants of cultivar Sabrina, according to Peterschmitt et al. (1996).

### Results

#### SP1 intra-isolate molecular variability

One hundred and twenty SP1 clones were obtained and analysed by RFLP. All but two clones showed the same restriction profile (Fig. 1a). R1, the clone previously obtained from isolate SP1 (Peterschmitt et al., 1996), belongs to this major group (group I). The group II clone (R10) was characterized by the absence of the BglII site in CP, and the group III clone (101) by a deletion in the LIR. Like R1, R10 was shown to be infectious by agroinfection.

Sequence analysis was carried out of the two minor-group clones, R1, and eight randomly selected group I clones. The sequence variations detected are shown in Table 1 and Fig. 2. The deletion detected by RFLP in clone 101 was shown to be located in an 87 bp region (positions 2652–2665) containing several deletions (38 nucleotides in total) and 28 mutations, which altered the stem–loop and conserved nonanucleotide sequences. Of the ten other clones which were sequenced not only in the LIR region (2255–2685), but also within CP and repA, six were identical and four had a single point mutation each, three of which (from clones 173, 195 and 237) were located in repA, and the other (from clone R10) in CP. When compared to R1 (identical to the SP1 consensus sequence; the R1 sequence was taken as the consensus in regions where only two sequences were available, see Fig. 2), the complete sequence of R10 revealed two more mutations (Table 1). Positions at which a single clone differed from the consensus are referred to as microheterogeneity (μ) sites. In total, six μ sites and a deletion in the LIR (clone 101) were detected in a total of 18240 nucleotides screened by sequencing.

#### SP2 intra-isolate molecular variability

After maintaining SP1 on a perennial host (C. lacryma-jobi) for 4 years, the resulting isolate, SP2, was investigated by RFLP analysis and sequencing. RFLP analysis of 36 SP2 clones showed that all but one (R13) had the same restriction profile as that from the major RFLP group (I) of SP1 (Fig. 1b). The RFLP variant, R13, was characterized by the absence of HindIII and EcoRV sites in repA.

Four SP2 clones were fully sequenced, three from the major RFLP group (R7, R11 and R12) and R13. Only R7 was tested for infectivity by agroinfection, and was shown to be

---

**Fig. 1.** Restriction maps of MSV clones obtained from isolates N2A, SP1 and SP2. Restriction sites are indicated above the diagrammatic representation of the 2685 nucleotide genome. The ORFs and the large and small intergenic regions (LIR and Sir, respectively) are indicated according to MSV-R (Peterschmitt et al., 1996), called here R1. Restriction sites not present in the various RFLP groups are represented by a horizontal bar. ∆ indicates a deletion detected in the LIR by RFLP analysis and shown by sequence analysis to involve 38 bp. Abbreviations for restriction sites are: AccI (Aii), AmII (AIII), BamHI (BHI), BglII (Bl), EcoRV (EV), HindIII (HIII), HindII (HII), HincII (HI), SacI (SI), Smal (Smi) and XhoI (XI). The BamHI restriction site was taken as position 1. The variants indicated for each group were sequenced fully except for 101, which was only sequenced partially.
Table 1. Mutations detected in SP1, SP2 and N2A clones

Mutations are described by comparison to the consensus sequence of each isolate. The last column identifies microheterogeneity sites (µ) and macroheterogeneity sites (M). Clones indicated by names starting with R have been sequenced fully; clones denoted by a number only have been sequenced partially. –, Not applicable.

<table>
<thead>
<tr>
<th>Isolate group</th>
<th>Clone</th>
<th>Nucleotides</th>
<th>Amino acids</th>
<th>Genomic region</th>
<th>Consensus</th>
<th>Variant</th>
<th>Position</th>
<th>Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) SP1</td>
<td>I 173</td>
<td>G T</td>
<td>repA</td>
<td>1778</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I 195</td>
<td>T C</td>
<td>repA</td>
<td>2313</td>
<td>Tyr</td>
<td>Cys</td>
<td>21 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I 237</td>
<td>T A</td>
<td>repA</td>
<td>2117</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II R10</td>
<td>C A</td>
<td>MP</td>
<td>200</td>
<td>Leu</td>
<td>Met</td>
<td>68 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G T</td>
<td>CP</td>
<td>456</td>
<td>Ala</td>
<td>Ser</td>
<td>48 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C G*</td>
<td>CP</td>
<td>722</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III 101</td>
<td>–</td>
<td>38 bp deletion*</td>
<td>–</td>
<td>LIR</td>
<td>– – –</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>(b) SP2</td>
<td>I R7</td>
<td>C G</td>
<td>CP</td>
<td>841</td>
<td>Ala</td>
<td>Gly</td>
<td>176 M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I R12</td>
<td>C T</td>
<td>repA</td>
<td>1976</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I 13</td>
<td>A C</td>
<td>repA</td>
<td>2573</td>
<td>LIR-R</td>
<td>–</td>
<td>– – –</td>
<td>µ</td>
</tr>
<tr>
<td></td>
<td>I 14</td>
<td>G C</td>
<td>CP</td>
<td>889</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G A</td>
<td>889</td>
<td>CP</td>
<td>989</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C A</td>
<td>1015</td>
<td>CP</td>
<td>2199</td>
<td>Leu</td>
<td>Trp</td>
<td>59 M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II R13</td>
<td>T G</td>
<td>MP</td>
<td>1526</td>
<td>Silent</td>
<td>–</td>
<td>µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T G</td>
<td>1525</td>
<td>repB</td>
<td>1600</td>
<td>Asp†</td>
<td>Ala</td>
<td>253 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T A</td>
<td>1640</td>
<td>repA, Arg</td>
<td>215</td>
<td>Ser</td>
<td>Thr†</td>
<td>215 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A C</td>
<td>1792</td>
<td>repA</td>
<td>2093</td>
<td>Ser</td>
<td>Ala</td>
<td>195 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G T</td>
<td>2093</td>
<td>repA</td>
<td>2155</td>
<td>Ser</td>
<td>Thr</td>
<td>74 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A T</td>
<td>2233</td>
<td>repA</td>
<td>2199</td>
<td>Ile</td>
<td>Leu</td>
<td>48 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T G</td>
<td>2233</td>
<td>repA</td>
<td>2217</td>
<td>Ala</td>
<td>Glu</td>
<td>53 µ</td>
<td></td>
</tr>
<tr>
<td>(c) N2A</td>
<td>I R2</td>
<td>T G</td>
<td>MP</td>
<td>35</td>
<td>Val</td>
<td>Leu</td>
<td>13 M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I 109</td>
<td>G T</td>
<td>CP</td>
<td>535</td>
<td>Glu</td>
<td>Asp</td>
<td>73 µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I 109</td>
<td>G A</td>
<td>CP</td>
<td>2466</td>
<td>Silent</td>
<td>–</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>II R3, R9</td>
<td>G T</td>
<td>repA</td>
<td>2093</td>
<td>Silent</td>
<td>–</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III R5</td>
<td>T G</td>
<td>SIR</td>
<td>1150</td>
<td>–</td>
<td>–</td>
<td>– – –</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>III R8</td>
<td>C G</td>
<td>repA</td>
<td>16444</td>
<td>Silent</td>
<td>–</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV R4</td>
<td>G T</td>
<td>repA</td>
<td>2217</td>
<td>Silent</td>
<td>–</td>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>

* Mutations responsible for RFLPs.
† Amino acids encoded by the spliced complementary sense transcript (Wright et al., 1997).
‡ Position 1664 is included in the complementary sense intron.

infectious. Six more clones were randomly selected and partially sequenced in CP, repA and LIR (Fig. 2).

The detected sequence variations are described in Table 1(b) and Fig. 2. Three clones, two that were sequenced partially and R11, had the same sequence as the SP2 consensus. The other seven clones contained one to seven mutations per clone. The mutations at positions 841 and 2199 were detected in three clones (R7, 3 and 21) and two clones (13 and 14),
Fig. 2. Intra-isolate heterogeneity of SP1, SP2 and N2A. The 2685 nucleotide consensus sequence of each isolate is represented diagrammatically, and ORFs and intergenic regions (LIR and SIR) are shown. The co-ordinates of the first and last nucleotides of the ORFs, as well as the approximate position of the stem–loop (2507–2554) present in the LIR, are given on the diagram below the three consensus sequences. The complementary sense intron is located between positions 1641 and 1732. The number of clones analysed in each region and for each isolate is given above each consensus, between horizontal arrows. The co-ordinates of the positions separating each of these regions are given at the top of the figure. The positions of nucleotide changes detected in a single clone are indicated by thin vertical lines (microheterogeneity sites, µ), whereas changes detected in more than one clone are indicated by thick lines (macroheterogeneity sites, M). All mutations are detailed in Table 1. The thick arrowhead under the SP1 consensus represents a region (2448–2534) where deletions and mutations were detected in the LIR of an SP1 clone (101). ★, positions where mutations led to RFLPs.

respectively. Positions at which a mutation was observed in more than one clone, but not in the majority, are referred to as macroheterogeneity (M) sites. In total, 16 µ sites and two M sites were detected in a total of 20070 nucleotides screened by sequencing.

N2A intra-isolate molecular variability

On the basis of RFLP analysis, the 40 N2A clones were divided into four groups, comprising 28 (group I), nine (II), two (III) and one (IV) clones (Fig. 1c). Each of the three smallest RFLP groups differed by only one restriction site from group I.

Seven N2A clones were fully sequenced (Fig. 2, Table 1c): two randomly selected clones from each of the two largest RFLP groups and all three remaining clones. Four of these clones differed by one nucleotide from the consensus sequence, two by two nucleotides and one by three nucleotides. All these clones, except R5 and R9, were shown to be infectious by agroinfection. Partial sequencing of the LIR and the beginning of repA was carried out on a further 15 N2A clones, 13 randomly selected clones from RFLP group I and two from group II. Eight group I clones were further analysed to provide sequences from within MP.

The number of positions at which nucleotide changes were observed was relatively low when compared with SP2 (Fig. 2, Table 1). For example, no variation was found in seven N2A sequences from the CP region, whereas five variable positions were detected in ten SP2 CP sequences. Likewise, only three variable positions were detected in seven N2A repA sequences, whereas seven were detected in ten SP2 sequences. However, the number of M sites was higher in N2A, which consists of at least four sub-populations: RFLP group I (28 clones), group II (nine clones), group III (two clones) and the R2-like group (four clones).

The four group II clones that were partially or fully sequenced were found to be identical. The only difference from the consensus was at position 2093, responsible for the RFLP. The two group III clones shared two M sites, one of which is responsible for the RFLP (position 1150). However they differed from each other at a µ site (Table 1c). R4, the only
M. Isnard and others

Fig. 3. Nucleotide differences between the consensus sequences of N2A, SP1 and SP2 isolates. (a) The long horizontal diagram represents the 2685 nucleotide MSV genome, with ORFs and intergenic regions (LIR, SIR) indicated. The shorter diagram is an enlargement of the region between positions 2255 and 2683. The region of the LIR to the left of the GC box and the stem–loop is labelled LIR-L, and the region to the right is LIR-R. The number of sequences used in each region to derive the consensus for each isolate is indicated above the diagrams. The co-ordinates of the positions separating each of these regions are given at the top of the figure. The positions of differences between the SP1 and SP2 consensus are represented by thick lines, whereas those between N2A consensus and the SP consensus (SP1, SP2) are represented by thin lines. (b) Details of the nucleotide differences between the SP and N2A consensus. (c) Details of the nucleotide differences between the SP1 and SP2 consensus. The columns labelled ‘Protein’ (b, c) indicate either the protein encoded by the gene containing the mutation, or the location of the mutation in the intergenic regions.

both differences lead to amino acid changes (Fig. 3c). One SP2 clone, R13, had the nucleotides found in SP1 clones at these two positions (Table 1b). All the N2A clones were identical to SP1 at these two positions.

Excluding positions 1640 and 2233, there are 11 differences between the N2A and SP (SP1 and SP2) consensus sequences (Fig. 3a, b). Eight differences were detected in intergenic regions, seven in the LIR and one in the small intergenic region (SIR). Three differences were found in ORFs: one non-synonymous mutation in repB and two silent mutations, one in repA and one in MP.

**Discussion**

**MSV isolates have quasispecies organization**

All three MSV isolates, SP1, SP2 and N2A, were shown to be heterogeneous by RFLP analysis and sequencing. As they
each consist of a mixture of genetically diverse but related virus genomes, they can be considered as quasispecies, according to Eigen et al. (1988). The mutations are distributed throughout the genome. The mutation frequencies (number of mutations relative to the consensus, divided by the number of nucleotides sequenced) are $3.8 \times 10^{-4}$ for SP1, $10.5 \times 10^{-4}$ for SP2 and $6.9 \times 10^{-4}$ for N2A. Interestingly, these mutation frequencies are in the same range as those obtained for plant RNA viruses: $10.2 \times 10^{-4}$ for tobacco mosaic virus (TMV) (Rodriguez-Cerezo & García-Arenal, 1989), $11.3 \times 10^{-4}$ for satellite TMV (Kurath et al., 1992) and $3.9-22 \times 10^{-4}$ for foreign sequences replicating in TMV (Kearney et al., 1993). Higher mutation frequencies, above $30 \times 10^{-4}$, are described for animal RNA viruses in portions of their genome reported to be variable (Adami et al., 1995; Lichtenstein et al., 1996). Since Drake (1991) showed that the mutation rate per base in DNA-based microbes is inversely proportional to genome size, it seems unsurprising that MSV heterogeneity was detected, since the geminivirus genome is one of the smallest among DNA viruses.

Intra-isolate heterogeneity has been investigated rarely for DNA viruses, probably because of their lower mutation rate compared to RNA viruses (Drake, 1991, 1993). Nevertheless, intra-isolate variability of a small sample of clones of cauliflower mosaic virus was investigated (Al-Kaff & Covey, 1994), and appeared to be higher than that of MSV, probably due to variations caused by its reverse transcription step (Chenault & Melcher, 1993). Also, it was shown by RFLP analysis that field isolates of BCTV can harbour mixed infections of two strains and, in some cases, contained more than one minor variant of the same strain (Stenger & McMahon, 1997).

**SP1 and SP2 isolates**

Based on RFLP analysis and partial sequencing, SP1 quasispecies appear to be relatively homogeneous when compared to N2A and SP2 (Figs 1 and 2, Table 1). As expected, it seems that SP1 selection through single-insect transmission operated as a genetic bottleneck. During its 5 min acquisition feed, the insect was obviously only able to penetrate a few cells, and it is unlikely that all the ingested virus could be transmitted during the 5 min inoculation feed (Reynaud & Peterschmitt, 1992).

Comparison of ten SP1 and ten SP2 clones in the LIR, CP and repA regions indicates that an increase in heterogeneity occurred during the 4 year maintenance of SP1 (Fig. 2): (i) 18 mutations were detected in SP2, but only four in SP1; (ii) six SP1 clones but only three SP2 clones had no mutations; (iii) there was only one mutation per clone for SP1, but often more than one in SP2 clones. A similar increase in molecular heterogeneity was observed after ten serial passages of infectious transcripts of a cucumber mosaic virus (CMV) satellite clone (Kurath & Palukaitis, 1990).

Another difference between SP1 and SP2 is the appearance of M sites corresponding to sub-populations in the SP2 quasispecies. This seems to be consistent with a recent model of RNA population evolution, which predicts that natural evolution of RNA populations is towards such sub-population organization (Huynen et al., 1996). Such an organization was also reported for satellite RNAs of CMV and TMV (Kurath & Palukaitis, 1989, 1990; Kurath et al., 1992; Kurath & Dodds, 1995).

The high heterogeneity of SP2 quasispecies is consistent with the hypothesis that virus quasispecies may show rapid evolution via accumulation of mutations when the population equilibrium is perturbed, in this case by host change and bottlenecks (Domingo et al., 1996).

The SP1 and SP2 consensus sequences differ by two non-synonymous mutations (Fig. 3a, c). These changes can be explained by selection of a minor variant in the SP1 population, either randomly through genetic bottlenecks and/or biologically through new cellular host conditions (Domingo & Holland, 1994). The contributions of the two events are generally difficult to estimate. Since large numbers of insects were used for virus transmission from maize to Coix and from Coix to maize, the risk of random selection is low at these steps. However, the selection of young sprouts from the old plant to maintain the isolate on the perennial host may have resulted in random selection. If a biological selection event had occurred, a minor variant might have been selected due to its higher fitness in the new host. The observation that both nucleotide substitutions encode amino acid changes in complementary sense proteins, which were shown for another geminivirus (WDV) to interact with host proteins (Xie et al., 1995), favours a potential role for biological selection. Such emergence of minor variants in virus quasispecies subjected to different environmental conditions has been studied with hepatitis C virus (Enomoto et al., 1994) and with human immunodeficiency virus type 1 (Wei et al., 1995).

**N2A isolate**

The number of heterogeneity sites detected in N2A was lower than that in SP2 (Fig. 2, Table 1). In spite of this relatively low molecular variability, it appears that N2A quasispecies are organized in sub-populations. Though such an organization would not be surprising in a quasispecies evolving in a tolerant environment, where new mutants have a high probability of being adapted and producing progeny, it is quite unexpected within an isolate selected in almost completely resistant hosts. It may indicate that the progressive increase in severity during N2A selection (see Methods) is concomitant with the selection of co-existing sub-populations maintained by complementation, as suggested for aphid-transmission mutants of a potyvirus population (Pirone & Blanc, 1996) and observed for a tobacco etch virus replication mutant (Li & Carrington, 1995). However, additional analysis will be needed to investigate this phenomenon further.
Comparison between N2A and SP1

Although the N2A and SP isolates originated from the same screening isolate, their consensus sequences differ at 11 positions, eight of which are located in the intergenic regions (Fig. 3a, b). These differences may be explained either by the initial heterogeneity of the screening isolate and the distinct selection procedures leading to N2A and SP1, or by genetic drift of the screening isolate during the 2 year gap between the selection of SP1 in 1991 and N2A in 1993. The ten differences observed between two SP2 clones (R12 and R13), despite the fact that SP2 was derived from a ‘biologically cloned’ isolate (SP1), suggest that the heterogeneity of the screening isolate might be high enough to produce two distinct isolates (N2A and SP1) differing at 11 positions. However, most of the differences between the SP isolates and N2A are concentrated in the LIR, whereas the differences between the clones within an isolate were distributed throughout the genome. This rather supports the drift hypothesis, since the LIR is known to be one of the most variable regions between MSV isolates from different geographical origins (Peterschmitt et al., 1996), suggesting that some of its sequences are submitted to relatively low selection pressures.

The authors wish to thank Dr Mustapha Bousalem for fruitful discussions. The work was supported by the European Commission (DGXI) and the French government.

References


Kurath, G., Rey, C. M. E. & Dodds, A. J. (1992). Analysis of genetic heterogeneity within the type strain of satellite tobacco mosaic virus reveals several variants and a strong bias for G to A substitution mutations. Virology 189, 233–244.


Quasispecies nature of maize streak virus


Received 15 April 1998; Accepted 5 August 1998