Presence of a polyA tail at the 3’ end of maize rayado fino virus RNA

Michael C. Edwards · John J. Weiland

Abstract Maize rayado fino virus (MRFV) is distinct from other marafiviruses in that its genome reportedly lacks a poly(A) tail at the 3’ terminus. We now show that the MRFV genome is indeed polyadenylated.

The family Tymoviridae is comprised of three genera: Tymovirus, Marafivirus, and Maculavirus [5]. Members of this family possess a 6–7.5-kb positive-sense RNA genome encapsidated within non-enveloped, icosahedral particles approximately 30 nm in diameter. The genome organization is similar among members of different genera, with methyltransferase, protease, helicase, and polymerase domains encoded sequentially in a large polyprotein open reading frame (ORF) encompassing most of the genome and coat proteins encoded near the 3’ terminus. Additional ORFs with a known or suspected role in movement or with other unknown function may be present. Tymovirus genomes typically possess a tRNA-like structure at the 3’ end, while both marafiviruses and maculaviruses are typically polyadenylated. Representatives of three of the four approved marafivirus species (Oat blue dwarf virus [7], Bermuda grass etched-line virus [10], and Citrus sudden death-associated virus [11]) and four tentative marafivirus species (Grapevine Syrah virus 1 [3, 14], Grapevine rupestris vein feathering virus [1], Grapevine asteroid mosaic-associated virus [1], Blackberry virus S [13]) are known to be polyadenylated. Poinsettia mosaic virus [4], an unassigned species in the family whose members share features with both the tymo- and marafiviruses, is polyadenylated, as is the recently identified olive latent virus 3 [2], a putative new member of the family Tymoviridae. Maize rayado fino virus (MRFV) is the type member of the genus Marafivirus, but it is distinct from other members of the genus in that its genome reportedly lacks a poly(A) tail at the 3’ terminus [9]. The purpose of this study was to reassess this apparent anomaly and confirm the nature of the MRFV 3’ terminus.

A Texas isolate of MRFV (MRFV-US; ATCC isolate PV-438) originally isolated from Zea mays by R.E. Gingery [8] was chosen for this study. The isolate induced characteristic MRFV symptoms in infected maize, reacted strongly in ELISA with MRFV-specific antiserum, and exhibited 98% and 95% sequence identity with the 3’ nontranslated region and partial CP sequence of the previously reported US and Costa Rican isolates of MRFV, respectively (not shown). These combined results confirmed the isolate’s identity and its close relationship to a fully sequenced Costa Rican isolate [9].

Two independent approaches were utilized to precisely determine the 3’ end sequence of the genomic RNA. In a naïve approach, cloning of ligation-anchored PCR (LA-PCR) [16] products provided the initial evidence that the 3’ end is polyadenylated. MRFV-specific clones were obtained by ligation of a dideoxy-terminated oligonucleotide (smartUcompdd; 5’ACTCTGCGTTGATACCACTGCTTGCCCTATAGTGAGTCGTATTAGddC3’) onto the 3’-end of the native MRFV-US RNA, followed by the use of a primer complementary to this (SmartNUP-A; 5’AAGCAGTGGTATCAACGCCAGT3’) for reverse transcription (RT) reaction priming and PCR amplification.
Briefly, a crude preparation of MRFV was produced according to Edwards and Weiland [6] from which RNA was extracted. Primer SmartUcompdd (50 pg) was mixed with 10 ng of RNA, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol, 20 units RNasin (Promega) and 20 units of T4 RNA ligase (New England Biolabs) in a volume of 10 μl, incubated at 37°C for 30 min, and heated to 65°C for 10 min. A portion (3 μl) of this ligated RNA was used as a target for RT-PCR in a 30-μl reaction containing 1X GoTaq Flexi Buffer (Promega), 1.5 mM MgCl₂, 0.1 mM each dNTPs, 30 ng each of the primers marafi5875fwd and smartNUP-A, 15 units of M-MLV reverse transcriptase (Promega), and 3 units of GoTaq DNA polymerase. The thermocycling parameters for the reverse transcription and DNA amplification were 50°C for 15 min, followed by 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. Reaction products were fractionated on a 5% polyacrylamide gel using Tris-borate EDTA running buffer. Products migrating between 400 and 600 bp were gel-purified according to standard methods [15], ligated into pGEM-T Easy (Promega), and the ligation was used to transform Z-competent E. coli DH5α cells (ZymoResearch). Clones were analyzed by restriction digestion and gel electrophoresis according to standard methods [15]; clones harboring inserts between 500 and 600 bp were subjected to sequence analysis.

Several LA-PCR-derived clones were chosen for sequencing. Three of these clones, designated pMRF18, pMRF20, and pMRF40, possessed a series of A residues ranging from 54 to 77 nucleotides in length at the 3' terminus (Fig. 1). Clones without the poly(A) sequence also were obtained, but these clones varied in length and terminated at random positions within the known MRFV sequence.

A second approach deliberately targeting putative poly(A) within the sequence was utilized to confirm these results. In this case, RT-PCR amplification was performed using native MRFV-US RNA as a template and an oligo(dT)₂₀⁻VN primer in both the priming of the RT reaction and in product amplification. This anchored primer, which allows priming only at the 5' end of a poly(A) sequence, was paired with primer marafi5875fwd for RT-PCR. MRFV RNA (10 ng) was mixed with 50 mM Tris-HCl (pH 8.3), 40 mM KCl, 8.75 mM MgCl₂, 10 mM DTT, 0.1 mg/ml acetylated BSA, 0.1 mM each dNTP, 0.4 ng oligo(dT)₂₀⁻VN primer, 40 units of RNasin (Promega), and 20 units of AMV Reverse Transcriptase in a 20-μl reaction volume. The mixture was incubated for 50 min at 42°C, and the reaction was terminated by heating to 94°C for 10 min. An aliquot (3 μl) of this RT reaction was used as template for PCR in a mixture containing 1X GoTaq Flexi Buffer (Promega), 1.5 mM MgCl₂, 0.1 mM each dNTP, 30 ng each of the primers marafi5875fwd and oligo(dT)₂₀⁻VN, and 3 units of GoTaq DNA polymerase. Thermocycling parameters were 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. As above, reaction products were fractionated on a 5% polyacrylamide gel, and a discrete product of ~380 nt, in agreement with the size predicted from the MRFV sequence, was purified from the gel. Products were cloned into pGEM-T Easy and analyzed as described above. Clones harboring inserts migrating at ~400 bp in agarose gel electrophoresis using TBE buffer were subjected to sequence analysis.

Fig. 1 Alignment between the 3' end of the genome of MRFV-Costa Rica (GenBank AF265566, nts 6283-6305) with clone sequences obtained from MRFV-US by RT-PCR. Clones including '3dT' in their names were obtained by RT-PCR from native MRFV-US RNA in which an oligo(dT)₂₀⁻VN primer served in both the priming of the RT reaction and in product amplification. Clones pMRF-18, pMRF-20, and pMRF-40 were obtained by ligation of a dideoxy-terminated oligonucleotide onto the end of the native MRFV-US RNA, followed by the use of a primer complementary to this for RT reaction priming and PCR amplification. The underlined ATC indicates the 3' half of the EcoRV cloning site of pGEM-T Easy, and the asterisks denote positions of complete conservation between the compared sequences.
Numerous clones possessing poly(A) tracts of 19-22 residues were obtained (Fig. 1). Both approaches clearly indicate the presence of a poly(A) tract at the 3′ terminus of MRFV-US. Furthermore, this poly(A) tract begins precisely at the end of the reported 3′-terminal sequence of MRFV from Costa Rica (GenBank no. AF265566) [9]. However, the consensus sequence at the junction of the poly(A) tail and the remainder of the genome includes two C residues, not a single C residue as previously reported for the 3′-terminal sequence. Whether the presence of two C residues reflects a sequence difference between the US and Costa Rican isolates remains to be determined.

Previous indications that the MRFV genomic RNA is not polyadenylated were based upon the reported inability to use oligo(dT) to prime cDNA synthesis [12] and unpublished attempts at direct sequencing of radiolabeled genomic RNA labeled at the 3′ end [9]. It is difficult to reconcile the results of these studies with those obtained in the present study. While the virus isolate used in our study is different than the Costa Rica isolate used in previous work, this alone does not provide a plausible explanation for the discrepancy. It may be that the MRFV genome 3′ end is particularly susceptible to degradation. Consistent with this, sequences of a number of clones in our study terminated at various nucleotide positions upstream of the 3′ terminus of the reported MRFV sequence. If polyadenylated RNAs represented a minority of the RNA population in previous studies, it is possible they would not be detected by the methods used in those studies, and any polyadenylation deliberately performed to facilitate the use of an oligo(dT) primer potentially would have masked the presence of a native polyadenylated sequence. Our methods, by contrast, were both PCR-based and as such would provide a more sensitive means of identifying full-length polyadenylated RNAs.

That the MRFV genome is polyadenylated is strongly supported by our strategy of using a sequence-independent method (LA-PCR) as a complement to a PCR-based method that allows priming only at the 5′ terminus of MRFV-US. Furthermore, this poly(A) tract begins precisely at the end of the reported 3′-terminal sequence of MRFV from Costa Rica (GenBank no. AF265566) [9]. However, the consensus sequence at the junction of the poly(A) tail and the remainder of the genome includes two C residues, not a single C residue as previously reported for the 3′-terminal sequence. Whether the presence of two C residues reflects a sequence difference between the US and Costa Rican isolates remains to be determined.

References

RNA polymerase in a newly identified species of plant alpha-like virus. Virology 394:1–7