Complete Cloning and Molecular Organization of a Rabies-related Virus,
Mokola Virus

By H. BOURHY,* N. TORDO, M. LAFON AND P. SUREAU
Unité de la Rage, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

(Accepted 13 April 1989)

SUMMARY

Mokola virus is a rabies-related virus responsible for both animal and human encephalitis cases in Africa. We report here the construction of a genomic library containing overlapping cDNA clones encompassing the entire genome. Five overlapping clones were sufficient to cover the genome (about 12 kb in size). Mokola virus was shown to share the same genomic organization as the rabies virus genome and also identical transcription signals. cDNA probes characterized six different transcripts by Northern blotting experiments; five of them corresponded in size and location to rabies virus mRNAs, and one of particular interest corresponded to a bicistronic RNA which included the genes for the phosphoprotein and the matrix protein. Comparison of the 3' and 5' end sequences of the Mokola genome with those of other members of the Lyssavirus genus showed a high homology and led us to propose a genus-specific consensus sequence. However, the latter appeared to vary widely throughout the Rhabdoviridae family.

INTRODUCTION

Rabies-related viruses have been isolated in different parts of the world (King & Crick, 1988) and classified into three serotypes on the basis of serological relationships, i.e. Lagos bat (serotype 2), Mokola (serotype 3) and Duvenhage (serotype 4). Together with the rabies virus (serotype 1), they constitute the Lyssavirus genus in the Rhabdoviridae family [Schneider et al., 1973; World Health Organization (1984) Technical Report Series, vol. 709]. Mokola virus was originally isolated in 1968 from pooled viscera of shrews (Crocidura sp.) captured in Ibadan, Nigeria (Shope et al., 1970; Kemp et al., 1972). Subsequent isolations were obtained from two Nigerian children with acute meningitis, one of whom died (Familusi & Moore, 1972; Familusi et al., 1972). Infected shrews were also reported in Cameroon (Le Gonidec et al., 1978). More recently, Mokola virus was isolated in Zimbabwe, during a widespread epizootic of domestic and wild animals (Foggin, 1982, 1983) and in the Central African Republic from a wild rodent (Lophuromys sp.) (Saluzzo et al., 1984). The wide distribution of Mokola virus in Africa suggests that infection of man and animal may occur more frequently than reported given the lack of epidemiological data concerning Africa. The origin of the reservoir of the virus is still unknown, but until now, shrews and rodents are the only reported wildlife host (Kemp et al., 1973). Although possible transmission through ticks or mosquitoes cannot be ruled out (Buckley, 1975; Aitken et al., 1984) shrew bite is probably the natural mode of transmission.

The reactivity pattern of different monoclonal antibodies directed against rabies virus and rabies-related viruses have outlined a distant relationship between Mokola and rabies viruses (Flamand et al., 1980a, b; Wiktor et al., 1984; Vincent et al., 1988; Bussereau et al., 1988). Unlike other rabies-related members (Lafon et al., 1986, 1988), Mokola virus is poorly neutralized by sera of rabies vaccinees (Wiktor et al., 1984; Celis et al., 1988) and poorly recognized by cytotoxic T lymphocytes specific for rabies virus (Wiktor et al., 1984, 1985; Cho et al., 1987; Celis et al., 1988). It has been shown in the laboratory and elsewhere that current potent rabies vaccines were consistently unable to protect mice against an intracerebrally inoculated Mokola virus challenge (Koprowski et al., 1985; unpublished observation). Such a
lack of cross-protection has been documented once in nature when a rabies virus-vaccinated dog died from encephalitis caused by Mokola virus (Foggin, 1982). This suggests that Mokola virus infection is possible before or after rabies virus vaccination and therefore is a potentially serious public health threat in Africa. The current lack of any efficient vaccine prompted us to undertake the molecular study of Mokola virus in order to increase understanding of its peculiar antigenicity and to provide a rationale for the genetic engineering of a future vaccine.

Mokola virus shares the structural organization of rabies virus, the prototype member of the Lyssavirus genus. The virion is bullet-shaped and contains an unsegmented and non-polyadenylated negative-stranded RNA genome (Sokol et al., 1969; Wagner, 1975). The transcription and replication mechanisms are thought to be similar to those of rabies virus and vesicular stomatitis virus (VSV) (Banerjee, 1987). The genome is sequentially transcribed into a plus-strand leader RNA, followed by five polyadenylated monocistronic mRNAs encoding successively the nucleoprotein (N), the phosphoprotein (M1), the matrix protein (M2), the glycoprotein (G) and the large protein (L). The latter acts as the major component of the viral RNA-dependent RNA polymerase (Ball & White, 1976; Flamand & Delagneau, 1978; Holloway & Obijeski, 1980; Iverson & Rose, 1981).

Because of the relatedness of the Mokola and rabies viruses, two rabies virus DNA primers were used to initiate the cloning of the Mokola virus genome. We describe the synthesis and characterization of cDNA clones covering the entire genome. A restriction map is included and the viral transcription products are analysed by Northern blotting experiments. We also present the 3' and 5' end sequences of the Mokola virus genome which are compared to that of the rabies virus.

**METHODS**

*Virus.* The strain studied was isolated from a cat in Zimbabwe (Foggin, 1982) and was analysed by Wiktor et al. (1984). This strain was first plaque-purified on CER cells according to the method of Wiktor & Clark (1973), and then propagated in BHK-21 cells and purified according to a modification of the technique described by Wiktor et al. (1977).

*Molecular cloning of Mokola virus.* Viral RNA was purified as described (Tordo et al., 1986a). The first strand cDNA synthesis was performed on 1 µg of genomic RNA using a 10-fold higher molar ratio of two different octadecamer primers (supplied by J. Igolen, Unité de Chimie Organique, Institut Pasteur, Paris, France) with nt positions 1 to 18 (3' primer) and nt positions 2901 to 2918 (M2 primer) (Tordo et al., 1985). The ds cDNAs were prepared according to the procedure of Gubler & Hoffman (1983) and separated by size using a Bio-Gel A-50m column (Huynh et al., 1985). The fractions (approx. 15 to 20 ng of cDNA) corresponding to the largest cDNAs were inserted in the PstI site of the pBR322 plasmid vector by the dC-/dG-tailing method (Rougeon & Mach, 1977). Competent *Escherichia coli* strain DH5-1 was transformed (Hanahan, 1983) and the transformants containing Mokola virus-specific inserts were detected by colony hybridization on nitrocellulose filters (Grunstein & Hogness, 1975).

*Mapping of cloned inserts.* Plasmid DNAs were isolated on a small scale by a modification of the alkaline extraction procedure (Birnboim & Doly, 1979; Maniatis et al., 1982). PstI inserts were electroeluted and purified by phenol–chloroform extraction and mapped by digestion with a range of restriction enzymes.

*Northern blot hybridization.* Cytoplasmic RNA extraction was performed as previously described (Ermine & Flamand, 1977). Total cellular RNA was denatured with 10 mM-sodium phosphate buffer pH 7.4, 50% formamide at 65 °C for 5 min and electrophoresed on 1 to 1.2% agarose gels containing 1:1 m-formaldehyde and 10 mM-sodium phosphate buffer pH 7.4. RNA was transferred and covalently fixed onto nylon membranes (Amersham). The cDNA probes were 3P-labelled using the multiprime DNA labelling system (Amersham) (specific activity, 106 c.p.m./µg). Conditions of prehybridization and hybridization were based on the method of Church & Gilbert (1984).

*DNA sequencing.* Restriction endonuclease cleavage fragments of the PstI inserts were subcloned into M13 vectors (Messing et al., 1981) and enzymic sequencing was performed by the chain-terminating inhibitor method (Sanger et al., 1977).

**RESULTS**

As our principal object was to provide a genetically engineered Mokola virus vaccine, we initiated cDNA synthesis using a rabies virus M2 primer assuming that it would specifically
Cloning of Mokola virus

2065

anneal to the Mokola virus genome. Therefore we expected to generate cDNA clones covering mainly the G-encoding gene located just downstream from the M2-encoding gene. The gene library contained approximately 5000 clones.

Detection of Mokola virus clones by rabies virus probes

Previous work on rabies virus had shown that, unlike the G-encoding genes which are very heterogeneous, the L and to a lesser extent the N species show the greatest degree of homology between the Vesiculovirus and Lyssavirus genera (Tordo et al., 1986a, b, 1988). Based on the premise that this homology should be even greater between members of the Lyssavirus genus, we investigated the ability of rabies virus 32P-labelled probes for N and L to cross-hybridize with Mokola virus mRNAs (Fig. 1). Even when these probes were located in some of the most conserved sequences of each gene encoding N and L, i.e. in nt positions 377 to 1039 (MboI-MboI) and nt 7034 to 7134 (HaeIII-HaeIII) respectively, a weak cross-hybridization with a transcript of the expected size was observed only with the rabies virus probe for N after a three-fold longer exposure. The lack of cross-hybridization within the L gene was at first surprising; however it could reflect the relative scarcity of the L mRNA in infected cells due to the polarity gradient of transcription (Iverson & Rose, 1981). Indeed the detection of the first Mokola virus-specific clone, pM12 (Fig. 2) with a probe covering nt 7901 to 9638 (HincII-HincII) of the rabies virus L gene confirmed the expected relatedness between the L genes of rabies Mokola viruses.

Characterization and isolation of cDNA clones representing the L gene and the 5' end of the Mokola virus genome

The location of the pM12 cDNA sequence in the L gene was further confirmed by the specific hybridization of a 60 bp HinfI-HinfI fragment to a large mRNA species among transcription products of Mokola virus-infected BHK-21 cells (Fig. 2, Fig. 4). Finally, partial sequence determination of the pM12 insert (unpublished data) demonstrated clear homology with the rabies virus L gene. We then selected two other Mokola virus-specific clones, pMB5 and pMR15a, that overlap with pM12 (Fig. 2), and determined the sequence of their ends. One end of the pMR15a insert appeared highly homologous with the 5' end of the rabies virus genome, and probably represented the 5' end of the Mokola virus genome (Fig. 3). Interestingly, the last 12 nucleotides were completely identical in the two genomes. Thus, the three Mokola virus-specific clones selected in the first cDNA library generated by the M2 primer, i.e. pMB5 (3300 bp), pM12 (2800 bp) and pMR15a (700 bp), covered the 5' half of the Mokola virus genome (Fig. 2). This indicates that the sequence of the M2 rabies virus primer might diverge from its counterpart on the Mokola virus genome, and was not adequate to initiate specifically the cDNA of the G gene. Therefore we decided to use a rabies virus 3' genomic end primer (nt 1 to 18) to complete the Mokola virus genome cloning, considering both the strong homology seen between the 5' ends of rabies and Mokola virus genomes, and the classical complementarity of 5' and 3' ends among unsegmented negative-stranded RNA virus genomes.

Completion of Mokola virus genome cloning

From the new colony library, two clones, pMD10 (4150 bp) and pMA10 (2850 bp), were selected by hybridization with previous clones and using restriction map analysis (Fig. 2). We sequenced the pMD10 end that was very homologous to the genomic 3' end of rabies virus (Fig. 3). We observed the structure of the rabies virus primer along the first 18 nucleotides and then a sequence strongly related to the rabies virus leader (+) gene. This proved that the priming of Mokola virus cDNA cloning with a 3' rabies virus primer was specific. Thus the entire Mokola virus genome was covered by five overlapping clones (ranging from the 3' end): pMD10, pMA10, pMB5, pM12 and pMR15a (Fig. 2). A detailed restriction map of the Mokola virus genome was determined using seven restriction enzymes, BamHI, BglII, EcoRI, HincII, HindIII, PstI and PvuII (Fig. 2).

Sequence analysis of the 3' and 5' ends of the Mokola virus genome

The structures of the 3' and 5' extremities of the Mokola virus genome are shown in Fig. 3 and are compared with corresponding regions in the PV strain of rabies virus. Two sequences
Fig. 1. Cross-hybridization of rabies virus probes located in the N and L genes with Northern blots of Mokola virus-infected cells. Total cytoplasmic RNAs from BHK-21 cells, taken at the indicated times after Mokola virus (lanes 1, 1a and 1b) or PV rabies virus (lanes 2) or mock (lanes 3) infections were electrophoresed on a 1.2% agarose-formaldehyde gel. Separated RNAs were blotted onto nylon membranes and probed with two 32P-labelled PV rabies virus genome-specific cDNAs at nt positions 377 to 1039 (MboI-MboI) [N probe; (a)] and nt 7034 to 7134 (HaeIII-HaeIII) [L probe; (b)]. A threefold overexposure is presented in lane 1a. Arrowheads indicate the position of the N and L rabies virus mRNAs.
Fig. 2. Restriction map and strategy for cloning the Mokola virus genome. The restriction map contains all the sites for the seven restriction enzymes. Arrows with black square ends correspond to *Bgl* II sites. The ends of the genome were sequenced and their structure is shown in Fig. 3. The Northern blot probes are defined as follows: N corresponds to the 1400 bp *PstI*-*EcoRI* fragment; N + M1 to the 1880 bp *HindIII*-*HindIII* fragment; M1 + M2 to the 2000 bp *EcoRI*-*EcoRI* fragment; G to the 2030 bp *FnuDI*-*SspI* fragment and L to the 60 bp *Hinfl*-*Hinfl* fragment. Sections of the genome are denoted by the proteins they encode.
Fig. 3. Comparison of the 3' and 5' ends of the Mokola virus genome with those of the PV strain of the rabies virus (Tordo et al., 1988). Complementary nucleotides of each virus are marked by long vertical lines. Consensus sequences specifying the N mRNA start, the L mRNA stop and the N protein start are indicated. The nucleotides of the genomic 3' and 5' ends of the Mokola virus that are identical to those of the PV strain are marked by short vertical lines. The residues are numbered from the genomic ends.
Fig. 4. Northern blot analysis of the mRNA species of Mokola virus. Ten μg of total cytoplasmic RNA from Mokola virus-infected BHK-21 cells, extracted 48 h post-infection, was electrophoresed through a 1% agarose gel. RNA blots were probed with N, N + M1, M1 + M2, G and L cDNA fragments of the Mokola virus genome (Fig. 2). Arrowheads indicate the position of 18S and 28S rRNA revealed by ethidium bromide staining and identify the different mRNA species revealed by autoradiography. No hybridization was apparent with total RNA from uninfected BHK-21 cells (not shown). The exposure time for each lane, which was optimized to give signals of similar strength, varied from 5 to 48 h.

Characterization of the Mokola virus mRNA species synthesized in vivo

We selected Mokola virus cDNA probes (Fig. 2) in order to characterize the in vivo transcription products during Mokola virus infection of BHK-21 cells. Six different transcripts were observed by Northern blotting experiments (Fig. 4). Faint bands appeared 6 h (not shown) and increased up to 48 h after infection. The characterization and the size of the Mokola virus

located at nt 59 to 67 from the 3' end (UUGUGAGGA) and nt 79 to 70 from the 5' end [AC(U)2G], are similar to the canonical rabies virus transcriptional start and stop signals respectively. We assume therefore that they represent the signal for initiation of transcription of the N gene and the signal for the termination of transcription for the L gene and for the polyadenylation of the L mRNA. At the 3' side of the N mRNA start signal and at the 5' side of the L mRNA stop signal, we observed two regions exhibiting no significant open reading frame. By analogy with the rabies virus genome, we suggest that they represent the leader (+) gene and the untranscribed 5' end respectively. They are of the same length as the rabies virus counterparts (58 and 69 nucleotides respectively), have an analogous heterogeneity in nucleotide composition (very rich in U and A respectively) and, more strikingly, show a high degree of homology with the rabies virus genome ends, noticeably increasing near the extremities and becoming identical for the 12 last nucleotides of the 5' end. Unfortunately, the use of the 3' rabies virus primer renders the 3' genomic ends relatedness between rabies and Mokola virus genomes impossible to evaluate by the method employed. However, the site-specific priming of the Mokola virus cDNA by the 3' rabies virus oligonucleotide strongly suggests that the 3' genomic ends of both viruses are as extensively homologous as their 5' end.

Characterization of the Mokola virus mRNA species synthesized in vivo

We selected Mokola virus cDNA probes (Fig. 2) in order to characterize the in vivo transcription products during Mokola virus infection of BHK-21 cells. Six different transcripts were observed by Northern blotting experiments (Fig. 4). Faint bands appeared 6 h (not shown) and increased up to 48 h after infection. The characterization and the size of the Mokola virus

located at nt 59 to 67 from the 3' end (UUGUGAGGA) and nt 79 to 70 from the 5' end [AC(U)2G], are similar to the canonical rabies virus transcriptional start and stop signals respectively. We assume therefore that they represent the signal for initiation of transcription of the N gene and the signal for the termination of transcription for the L gene and for the polyadenylation of the L mRNA. At the 3' side of the N mRNA start signal and at the 5' side of the L mRNA stop signal, we observed two regions exhibiting no significant open reading frame. By analogy with the rabies virus genome, we suggest that they represent the leader (+) gene and the untranscribed 5' end respectively. They are of the same length as the rabies virus counterparts (58 and 69 nucleotides respectively), have an analogous heterogeneity in nucleotide composition (very rich in U and A respectively) and, more strikingly, show a high degree of homology with the rabies virus genome ends, noticeably increasing near the extremities and becoming identical for the 12 last nucleotides of the 5' end. Unfortunately, the use of the 3' rabies virus primer renders the 3' end relatedness between rabies and Mokola virus genomes impossible to evaluate by the method employed. However, the site-specific priming of the Mokola virus cDNA by the 3' rabies virus oligonucleotide strongly suggests that the 3' genomic ends of both viruses are as extensively homologous as their 5' end.

Characterization of the Mokola virus mRNA species synthesized in vivo

We selected Mokola virus cDNA probes (Fig. 2) in order to characterize the in vivo transcription products during Mokola virus infection of BHK-21 cells. Six different transcripts were observed by Northern blotting experiments (Fig. 4). Faint bands appeared 6 h (not shown) and increased up to 48 h after infection. The characterization and the size of the Mokola virus
transcripts, revealed by successive cDNA probes, suggests that the Mokola virus transcriptional map is identical to that of rabies virus (Flamand & Delagneau, 1978). Indeed five monocistronic mRNA species successively appeared from the genomic 3' end. The lengths of four of these RNAs were estimated on agarose gels to be 1750, 1250, 1000 and 2400 nt including the poly(A) tail. They are thought to encode N, M1, M2 and G respectively. Only the size of the fifth one, the L mRNA, was lower (4800 nt) than expected from the genomic length (6500 nt). Curiously, a sixth transcriptional product appeared in lane N + M1 and M1 + M2 of Fig. 4 corresponding to a bicistronic M1-M2 mRNA as suggested by its Mr, corresponding to 2200 nt. By comparison, other Mokola virus polytranscripts occurred at low frequency in infected BHK-21 cells and were seen only after overexposure (not shown). It is also noteworthy that we failed to characterize the equivalent of the leader RNAs found during the transcription of rabies virus (Kurilla et al., 1984) or VSV (Colonno & Banerjee, 1978).

**DISCUSSION**

The RNA polymerase of RNA viruses is well known to show a very high error frequency resulting in a wide distribution of mutations throughout the viral population (Domingo et al., 1985; Steinhauer & Holland, 1986). Therefore we decided to plaque-purify the Mokola virus in order to study a relatively homogeneous subpopulation. We selected one Mokola virus strain giving a very large c.p.e. in CER cells, confirming the plasticity of this phenotypic character throughout the population (Clark & Wiktor, 1974). This strain gave sufficient yields for virus purification and exhibited a high pathogenicity in adult mice (3 p.f.u./LD50 administered intracerebrally). This indicated that the adaptation to BHK-21 cells and plaque purification did not reduce the pathogenicity of the strain studied.

The construction of a recombinant cDNA library extending continuously from the 3' end to the 5' end of the Mokola virus genome is reported. Here we show five cDNA clones covering the entire genome with minimal overlap. The sequences of the expected 3' and 5' ends were determined and they exhibit a high level of homology with those of the rabies virus genome. The genome of Mokola virus is an unsegmented negative strand RNA whose size, estimated at 12800 nt is similar to that of the rabies virus genome (11932 nt) (Tordo et al., 1988). The characterization of the N mRNA start sequence [AACACTCCT, in the antigenomic (+) sense] and the L mRNA stop sequence [TG(A)TC] which together match the rabies virus canonical transcription sequences (Tordo et al., 1986b), suggested that both viruses share similar regulatory mechanisms.

The 18 terminal nucleotides of the 3' end of the Mokola virus genome remain to be determined because a rabies virus primer was used to initiate Mokola virus cDNA. However, two independently derived sets of data strongly suggest that the 3' and 5' ends of the Mokola virus genome are complementary. Firstly, the 12 5' end nucleotides of the genome are strictly identical to those of the PV strain of rabies virus, in which 11 are complementary to the homologous region at the 3' end (Tordo et al., 1988). Secondly, specific 3' priming of the Mokola virus cDNA by a 3' rabies virus primer reveals a close homology between both virus 3' ends. The terminal complementarity, consistently found throughout unsegmented negative-stranded RNA viruses sequenced to date (Keene et al., 1979; Shioda et al., 1986; Nichol & Holland, 1987; Tordo et al., 1988; Crowley et al., 1988) is unlikely to give rise to a panhandle structure in the genomic RNA because this RNA is encapsidated as replication proceeds (Chanda & Banerjee, 1979). It may however reflect the conservation of signals thought to be present at the genomic ends such as the initiation site of RNA synthesis and the nucleation site for the initiation of nucleocapsid assembly (Banerjee, 1987).

Previous comparative studies of PV and CVS rabies virus strains, members of serotype 1 of the Lyssavirus genus, have shown that mismatches between 3' genomic ends do not appear before nucleotide 13 (Tordo et al., 1988; Poch et al., 1988). The present study extending this observation to another serotypic group, indicates that lyssaviruses may have identical genomic 5' ends up to position 11: 3' UGCCAAAUGGUU (genomic sense) (Fig. 5). This proposal will be confirmed by sequencing members of the two remaining serotypes of the Lyssavirus genus. At
Cloning of Mokola virus

Consensus sequence of 3' genomic and antigenomic ends

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>UGC</td>
<td>GAA</td>
<td>UGGU- -URUP-</td>
</tr>
<tr>
<td>UGC</td>
<td>RFU</td>
<td>-U- -UUGGU</td>
</tr>
</tbody>
</table>

Genus               | Family            |
---------------------|-------------------|
Lyssavirus           | Rhabdoviridae     |
Vesiculovirus        |                   |

Fig. 5. Comparison of the invariant sequences at the 3' genomic and antigenomic ends of the Lyssavirus genus members with those of the Vesiculovirus genus and those of the Paramyxoviridae family. P and Py represent purine and pyrimidine residues respectively. The sequences are drawn from Giorgi et al. (1983), Shioda et al. (1983, 1986), Blumberg et al. (1984), Roy et al. (1984), Kurilla et al. (1984), Dimock et al. (1986), Tordo et al. (1986a, b, 1988), Nichol & Holland (1987), Sakai et al. (1987), Yusoff et al. (1987), Crowley et al. (1988) and Poch et al. (1988).

This seems very likely because the two studied serotypes, 1 and 3, appear the most divergent by monoclonal antibody analysis (Flamand et al., 1980a, b; Wiktor et al., 1984). In the Vesiculovirus genus, an analogous conserved terminal octadecanucleotide was characterized by the comparison of five different strains (Nichol & Holland, 1987). However when the consensus sequences of both genera are compared, there is little similarity, except at the identical initial codon UGC (Fig. 5). This indicates that the conserved terminal sequence is typical for each genus and diverges rapidly according to evolutionary distance. However viruses having this class of genome shared a characteristic nucleotide composition which is very rich in U residues at the genomic 3' end. This feature, particularly noticeable between nt 7 and 40 (Fig. 3) was also observed in lyssaviruses of serotype 1 (Tordo et al., 1988; Poch et al., 1988). This region may include the site of interaction with the M1 protein which is an important cofactor of the transcription complex (Keene et al., 1981; Isaac & Keene, 1982).

This paper also characterizes the major RNA transcripts produced during Mokola virus infection of BHK-21 cells and demonstrates that the gene order is identical to that of rabies virus. Hybridization of ds cDNA probes, progressively covering the entire Mokola virus genome, to blots of infected cytoplasmic total RNA revealed five transcripts corresponding both in size and location to those of rabies virus (Tordo & Poch, 1988). We assume therefore that the Mokola virus genome is transcribed into five monocistronic mRNAs encoding, successively from the 3' end, N, M1, M2, G and L. There is a good correlation between the protein-encoding potential of each transcript, as deduced from their electrophoretic mobility, and the measured Mr values of the proteins themselves (Sokol & Koprowski, 1975; Coslett et al., 1980). Our failure to detect plus and minus leader RNAs by Northern blot experiments suggests that their transcription could occur at a level significantly lower than in VSV (Leppert et al., 1979) as was the case for measles virus (Crowley et al., 1988). Whatever the cause, the detection of such small transcripts requires techniques more appropriate than Northern blotting.

The most striking feature of Mokola virus transcription is that the M1- and M2-encoding genes exhibit a particular tendency to be tandemly transcribed as an M1–M2 bicistronic transcript. An approximate quantitative analysis indicates that this transcript occurs as frequently as monocistronic M1 and M2 transcripts. Polytranscripts are normally rare products of infection by unsegmented negative-stranded RNA viruses and are thought to be aberrant transcripts generated by occasional termination failures in the stop–start mechanism of
transcription (Banerjee, 1987). Such failures appear to occur more frequently in the intergenic region linking the phosphoprotein and the matrix protein genes. In VSV, a 10% increase in readthrough frequency has been noted at this junction and is presumed to be due to a single transition, G to C, in the dinucleotide composing the intergenic sequence (Masters & Samuel, 1984). Furthermore the matrix protein monocistronic mRNA usually produced during transcription of measles virus was completely replaced by a bicistronic mRNA containing both the phosphoprotein and matrix protein gene sequence in a subacute sclerosing panencephalitis brain (Cattaneo et al., 1986). Whether the characteristic phosphoprotein and matrix protein bicistronic RNA plays a particular role in the biology of these viruses remains unclear.

However, it is noteworthy that up to now within the Lyssavirus genus the M1–M2 bicistronic transcript appears to be characteristic of only the Mokola virus and does not exist as extensively during rabies virus transcription in BHK-21 cells (Holloway & Obijeski, 1980; Tordo & Poch, 1988). The sequence determination of the Mokola virus M1–M2 intergenic regions, currently in progress, will be of great interest in providing a molecular basis to this discrepancy.

Mokola virus studies which were initially intended to help the development of a specific vaccine have been shown to provide important fundamental information about the genomic structure of unsegmented negative-stranded RNA viruses.

We are indebted to Professor F. Rougeon and his group (Unité de Génétique et Biochimie du Développement–Institut Pasteur) who kindly provided facilities during this project. We are grateful to Michele Storrs for critical reading of the manuscript. This work was supported by grants from the Comité Consultatif des Applications de la Recherche de l’Institut Pasteur (7830).

REFERENCES


Cloning of Mokola virus

Sequence variability and function of measles virus 3' and 5' ends and intercistronic regions. Virology 164, 498–506.


(Received 16 December 1988)