Evidence for recent genetic variation in monkeypox viruses

N. J. Douglass, M. Richardson and K. R. Dumbell*

Department of Medical Microbiology, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, South Africa

DNA from isolates of monkeypox virus, when digested with the endonuclease \textit{PstI}, gave fragment-size profiles which correlated with the geographic area from which the isolate originated. Although some of the differences were located subterminally in the genome, others mapped to the central conserved region. Further differentiation of the viral genomes was sought by analysis of a short region within the central conserved part of the genome that appeared to be a partially deleted counterpart of an intact 1024 bp open reading frame (ORF) present in variola and vaccinia virus genomes. We reasoned that this region would not be conserved by functional selection and would therefore be likely to show more variation between isolates of monkeypox virus. The deletions found in monkeypox virus isolates from Liberia and from Benin were almost the same as that which we had previously found in the Denmark strain. A much shortened ORF, potentially coding for a product of 133 amino acids, was retained in all three West African isolates, but three Zairean isolates each showed an identical series of small insertions and deletions which effectively abolish the ORF. Three deletions, present in all isolates, must pre-date the geographical separation of monkeypox virus lineages; other, presumably more recent, changes differ between the Zairean and West African isolates. In contrast, the base similarity was found to be more than 99\% when all the monkeypox virus sequences were appropriately aligned. This, in a disrupted and presumably non-functional gene also indicates that the changes described are recent. It is suggested that insertions and deletions occur regularly during poxvirus DNA replication, but are preserved only in sequences that are not required for continued transmission in the natural host.

Introduction

Monkeypox virus was first isolated in Denmark in 1958 (Von Magnus \textit{et al}., 1959) from captive monkeys imported from Singapore. Later outbreaks occurred in Europe and the United States affecting a range of animals; the index case originated from either Asia or West Africa (Jezek & Fenner, 1988). To locate the natural reservoir of monkeypox virus the WHO conducted extensive serological surveys in Asia and Africa. Initially, no \textit{Orthopoxvirus} antibodies were detected in sera from either continent (Arita \textit{et al}., 1972).

In 1970 the first case of human monkeypox was recognized in Zaire (Ladnyj \textit{et al}., 1972). Further outbreaks were reported in countries of West and Central Africa, occurring infrequently in isolated villages within the tropical rain forests. Between 1970 and 1986 human monkeypox was detected in seven African countries (Jezek \textit{et al}., 1988) and more recently a case was described in an eighth, Gabon (WHO, 1992). Attention has therefore been focused on West and Central Africa, which is the only region in which natural monkeypox virus has been found.

\textit{Orthopoxvirus} antibodies have been detected in sera from West African primates (Breman \textit{et al}., 1977), three of which contained monkeypox virus-specific antibodies (Gispen \textit{et al}., 1976). Further serological investigations in Zaire (Marennikova \textit{et al}., 1975; J. H. Nakano, quoted in Jezek & Fenner, 1988) revealed monkeypox virus-specific antibodies in 10 species of monkey from four genera, and four species of tree squirrel from two genera. In addition monkeypox virus has been recovered from a diseased squirrel from the agricultural area adjoining a village in Zaire (Khodakevich \textit{et al}., 1986; Marennikova \textit{et al}., 1986). It is likely that squirrels, which have a high population density, are the natural host for monkeypox virus (Khodakevich \textit{et al}., 1987), and that humans and monkeys are incidental hosts.

The DNA sequence data presented in this paper have been submitted to GenBank and have been assigned the following accession numbers: VVDIE, M95534; MPDENMAR, M95534; MPLIBER, L20440; MPBENIN, L20439; MPZ8621, L20436; MPZ86112, L20437; MPZ1324, L20435; MPZ241, L20438.
naturally occurring monkeypox virus in any other part of the world.

The genomes of 20 monkeypox virus isolates (and 18 other Orthopoxviruses) have been mapped for *Hind*III cleavage sites (Esposito & Knight, 1985). All monkeypox viruses have the same pattern in the central conserved part of the genome, but differences are found at the termini. These differences do not differentiate human from animal isolates but instead appear to group the viruses according to geographical location. The West African isolates form two groups, those from the Liberia–Sierra Leone region and those from Nigeria. The Zairean isolates form the third group. Of the four Zairean isolates, there is one variant with an extra *Hind*III site near its left end. Here we report the *Pst*I profiles and maps of the three different monkeypox virus groups: West Africa, Nigeria and Zaire. Differences were found within the conserved region of the genome as well as the termini.

Monkeypox virus DNA sequence data are limited to short stretches from the Denmark strain only (Esposito & Knight, 1984; Cowley & Greenaway, 1990; Douglass & Dumbell, 1992). These sequences have been compared with those of variola and vaccinia viruses, but there has been no reported comparison of DNA sequence within the monkeypox virus genome. We have sequenced the region equivalent to the *E5R* open reading frame (ORF) in the Copenhagen strain of vaccinia virus (Goebel et al., 1990) in five monkeypox virus isolates from different parts of West and Central Africa. The DNA sequences were compared to those from monkeypox virus Denmark and variola virus Harvey (Douglass & Dumbell, 1992). Both inter- and intra-species differences were noted and are discussed.

**Methods**

**Viruses.** The authors thank Dr S. S. Marennikova and the late Dr J. H. Nakano for providing monkeypox virus isolates from human infections in Africa, Dr R. Gispen for the Rotterdam strain and the late Dr J. H. Nakano for the squirrel isolate. Other isolates were provided by arrangement with the WHO.

**Preparation of viral DNA.** Genomic viral DNA was prepared as described by Dumbell & Richardson (1993).

**Restriction enzyme digestion.** Restriction enzymes were purchased from Amersham and Boehringer Mannheim. Digestions were performed according to standard methods (Mantatig et al., 1982) and the DNA fragments were separated by electrophoresis in 0.8% agarose gels at approximately 2 V/cm.

**DNA mapping.** Monkeypox virus *Pst*I maps were constructed relative to the known monkeypox virus Denmark *Hind*III map (Mackett & Archard, 1979) by cross-hybridization experiments. Initially *Hind*III clones of monkeypox virus Denmark were radiolabelled and used to probe *Pst*I digests of monkeypox viruses. Differences between the three monkeypox viruses were clarified by cross-hybridization between the monkeypox DNA blots and labelled *Pst*I fragments eluted from gels.

Digested DNA was transferred to nylon membranes (Amersham) by Southern blotting (Southern, 1975). DNA fragments were eluted by the freeze-squeeze method (Rickwood & Hames, 1982), and radiolabelling was achieved by nick translation (Maniatis et al., 1982). Amplification was performed as described by Johnson et al. (1984).

**Amplification of monkeypox virus DNA equivalent to vaccinia virus E5R.** PCR amplification of DNA was performed according to standard methods on a Techne thermocycling machine using Taq DNA polymerase supplied by BRL. Oligonucleotide primers were designed to bind to DNA sequences flanking the vaccinia virus *E5R* gene. The primer sequences were: 1: *5' GATGATTTTCCAGGCCCATT 3'* and 2: *5' GAGCTAGTACATGATTGAGGGT 3'*.

An annealing temperature of 50 °C was used for 1 min and extension times were either 3 min or 1.5 min at 72 °C.

**Amplification of the DNA flanking the major deletion in monkeypox virus.** Primer sequences used for differentiating monkeypox viruses into two groups were 1: *5' CCCAATTGTGGAGGAGG 3'* and 2: *5' CGA
digested with *Pst*I for one example of each of the main groups. Monkeypox viruses Denmark, Benin and Z85-249 were chosen for this purpose, and the resulting maps are shown in Fig. 2.

**Cloning of the E5R-equivalent region of monkeypox virus DNA.** PCR-amplified DNA fragments were ligated into the *Smal* site of D-tailed pUC18 (Marchuk et al., 1991). Competent LKIII cells were transformed (Chung & Miller, 1988) and positive clones were detected by standard methods for colony blot hybridizations and restriction enzyme analysis (Maniatis et al., 1982).

**DNA sequencing.** Double-stranded DNA was sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977) using [*32P]*dATP (Amersham) and Sequenase kits (United States Biochemical). Template DNA was prepared according to a modified method of Birnboim & Doly (1979). Forward and reverse M13 primers were used as well as other synthesized oligonucleotides.

**Results**

**Cleavage site maps**

*Pst*I restriction enzyme digests of monkeypox virus DNA gave three main fragment size profiles as shown in Fig. 1. One pattern was common to three isolates from Liberia and Sierra Leone together with six isolates from captive animals in Europe and America, as listed in Table 1. A second profile was given by two isolates from Nigeria and the third was typical of seven isolates from Zaire. (Minor variations affecting the size of some of the fragments were seen within the group of Zairean isolates.)

The *Pst*I cleavage sites were mapped on the genome for one example of each of the main groups. Monkeypox viruses Denmark, Benin and Z85-249 were chosen for this purpose, and the resulting maps are shown in Fig. 2. Esposito & Knight (1985) noted that *Hind*III maps of monkeypox virus DNA distinguished between isolates from West Africa and Zaire on the basis of variation in the size of fragments mapping near the left-hand
Table 1. Details of monkeypox virus isolates

<table>
<thead>
<tr>
<th>Isolate designation</th>
<th>Year of isolation</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>1958</td>
<td>Captive monkey from Singapore</td>
</tr>
<tr>
<td>Prier</td>
<td>1959</td>
<td>Captive monkey from Malaysia</td>
</tr>
<tr>
<td>McConnell</td>
<td>1962</td>
<td>Captive monkey, Washington D.C., U.S.A.</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>1964–1965</td>
<td>Captive animals, Rotterdam Zoo, Netherlands</td>
</tr>
<tr>
<td>Espana</td>
<td>1967</td>
<td>Captive monkeys from India, Malaysia</td>
</tr>
<tr>
<td>Paris</td>
<td>1968</td>
<td>Captive chimpanzee from Sierra Leone</td>
</tr>
<tr>
<td>Liberia I</td>
<td>1970</td>
<td>Human, Liberia</td>
</tr>
<tr>
<td>Liberia II</td>
<td>1970</td>
<td>Human, Liberia</td>
</tr>
<tr>
<td>MP-266</td>
<td>1970</td>
<td>Human, Sierra Leone</td>
</tr>
<tr>
<td>MP-82</td>
<td>1971</td>
<td>Human, Nigeria</td>
</tr>
<tr>
<td>Benin</td>
<td>1978</td>
<td>Human, Nigeria</td>
</tr>
<tr>
<td>Z241</td>
<td>1972</td>
<td>Human, Katakoko-kombe, Kasai, Zaire</td>
</tr>
<tr>
<td>Z369</td>
<td>1973</td>
<td>Human, Lisala, Equateur, Zaire</td>
</tr>
<tr>
<td>Z1324</td>
<td>1979</td>
<td>Human, Gemena, Equateur, Zaire</td>
</tr>
<tr>
<td>85-249</td>
<td>1985</td>
<td>Squirrel, Bumba, Equateur, Zaire</td>
</tr>
<tr>
<td>85-240</td>
<td>1985</td>
<td>Human, Bumba, Equateur, Zaire</td>
</tr>
<tr>
<td>86-21</td>
<td>1986</td>
<td>Human, Bumba, Equateur, Zaire</td>
</tr>
<tr>
<td>86-112</td>
<td>1986</td>
<td>Human, Bongadanga, Equateur, Zaire</td>
</tr>
</tbody>
</table>

within the Benin and Denmark PstI K fragment. In the region of HindIII D (Denmark), the PstI site separating the L and M (Denmark) or K and L fragments (Zaire) is missing in the Benin isolate. Within the region HindIII A (Denmark), the PstI site separating Zaire F and H is not present in Denmark or Benin isolates.

DNA sequence

In a previous comparison of the genomes of monkeypox and variola viruses, an ORF conserved in variola and vaccinia viruses was found to be truncated in monkeypox virus Denmark (Douglass & Dumbell, 1992). This ORF is equivalent to the vaccinia virus Copenhagen ES5 ORF. The equivalent region was sequenced in five African monkeypox virus isolates and compared to the variola virus Harvey and monkeypox virus Denmark sequences.

Fig. 3 shows the alignment of the DNA sequences

---

Fig. 1. Agarose gel electrophoresis of digested monkeypox virus DNA. Lane 3 is a HindIII digest of monkeypox virus Denmark; lanes 4 to 6 are PstI digests of monkeypox viruses Denmark (West Africa), Z85-249 (Zaire) and Benin (Nigeria) respectively. Lanes 2 and 7 are HindIII digests of lambda phage DNA and lanes 1 and 8 are undigested lambda phage DNA.

terminus. The PstI maps show this feature but also show cleavage site differences in the central conserved region.

In the region of HindIII C (Denmark), the Zaire PstI P and Q fragments are formed by an extra cleavage site
Fig. 3. For legend see opposite.
starting at a common NcoI site and ending at the variola virus termination codon. The E5R ORF of vaccinia virus Copenhagen starts at position 207 and ends at position 1214. An insertion of 2 bp in variola virus results in an ORF which starts at position 143 and ends at position 1214.

There are a number of base substitutions and deletions in the corresponding monkeypox virus sequences. Four of these differences alter the ORF and are found in all the monkeypox virus isolates. At position 263 a base transversion (G to T) generates a stop codon in frame.

Two 24 bp deletions are present at positions 200 to 223 and 1154 to 1177. These sequences are flanked by short direct repeats in variola virus; only one copy of the repeat sequence is present in monkeypox virus. The first deletion is associated with a 6 or 7 bp direct repeat, and the second with a 6 bp direct repeat. These deletions may have arisen by the 'slipped mispairing model' (Albertini et al., 1982; Levinson & Gutman, 1987). Similar deletions between short repeats have been noted previously in vaccinia and variola viruses (Smith et al., 1991; Aguado et al., 1992).

The fourth difference common to all monkeypox virus sequences is a single base pair deletion at position 603. The large deletion of 391 bp that was demonstrated (Douglass & Dumbell, 1992) in the Denmark strain was also found in all the other West African isolates. In the corresponding region in the Zaire isolates there are three deletions of 5 bp (positions 697 to 701), 9 bp (positions 815 to 823) and 127 bp (positions 867 to 993) as well as an insertion of 1 bp at position 1003.

Common to the Zaire isolates only are additional deletions of 1 and 2 bp and an insertion of 3 bp (positions 308, 432 to 433 and 512 to 514, respectively).

The differences found differentiate monkeypox virus into two groups: West Africa (isolates Benin, Liberia II and Denmark) and Zaire (isolates 86-21, 86-112 and Z1324).

The size of the large deletion was investigated in other monkeypox virus isolates by PCR. DNA spanning the deletion was amplified to generate a fragment of either 342 bp (Zaire isolates) or 91 bp (West African isolates). Consistent with the sequencing data, all isolates fell into one of two groups depending on the geographical origin of the virus (Fig. 4). All six viruses from Zaire gave an amplified fragment of 342 bp, whereas five from countries further west gave a fragment size of 91 bp. Each of the four isolates from captive animals resembled the West African monkeypox virus. This is additional evidence that the source of virus for the outbreaks in Europe and the United States was West Africa. Variola and vaccinia viruses gave a fragment size of 482 bp as did isolates of rabbitpox virus, buffalopox virus and camelpox virus.

All three Zairean isolates have identical sequences. Also, the sequence of the Liberia II isolate is identical to that of the Denmark strain, suggesting that the Denmark strain originated in this region. With the alignment shown in Fig. 3, all seven monkeypox virus isolates share more than 99% base similarity.

**Discussion**

It has been previously noted (Esposito & Knight, 1985) that monkeypox virus HindIII restriction profiles are characteristic for viruses from West Africa, Nigeria or Zaire. In the HindIII profiles the differences are in the size of fragments near the termini, but the differences in
Fig. 4. Agarose gel electrophoresis of PCR-generated DNA fragments corresponding to the region of the major deletion in monkeypox virus. Lanes 2 to 4: vaccinia viruses Dairen, rabbitpox and buffalopox, respectively; lanes 5 to 6: camelpox viruses 79-241 and 2740, respectively; lanes 8 to 12: monkeypox viruses, Zaire isolates 85-249, 86-112, 85-240, 1024 and 1324, respectively; lanes 14 to 16: monkeypox viruses, West African isolates Liberia I, MP-82 and MP-266 respectively; lanes 17 to 19: monkeypox viruses from captive animals, isolates Prier, Espana and Paris, respectively; lane 21: negative control (no DNA); lanes 1, 7, 13 and 20 contain DNA markers (Boehringer Mannheim marker VI).

Fig. 5. Diagram to show the relationships of the monkeypox viruses from the seven base mismatches found within the aligned 767 bp.

the PstI profiles map to the conserved region of the genome as well as to the terminal regions.

A gene sequence within the central conserved region, which previous work (Douglass & Dumbell, 1992) had shown to be degenerate, was analysed in more detail in the belief that mutational changes would accumulate more rapidly there than in sequences subject to functional selection.

All six monkeypox virus isolates that were sequenced had a single base mutation introducing a stop codon (in frame with the ATG at position 185) at position 263 and three deletions of 24 bp, 1 bp and 24 bp. These changes can be presumed to have been present in an ancestral monkeypox virus; other changes, which are presumably more recent, have occurred independently in the monkeypox virus lineages in West Africa and Zaire and differentiate these two groups. When the monkeypox virus sequences are aligned, as in Fig. 3, the seven positions of mismatched bases are grouped, according to the geographic origin of the isolates as shown in Fig. 5.

It was noted that the location of the two 24 bp deletions were flanked in the variola virus sequence by short direct repeats, although for the deletion at positions 200 to 223 the left-hand repeat in the variola virus sequence has only six of the seven bases in the monkeypox virus equivalent sequence. Smith et al. (1991) and Aguado et al. (1992) suggested that deletions which occur in a region flanked by short direct repeats may be a possible mechanism for the evolution of poxvirus genomes. It was noted that of the various deletions in the monkeypox virus genome it was only the two most ancient that showed this feature.

The current Zairean isolates have no ORF of meaningful length in the region corresponding to the variola virus ORF. In the West African isolates there is a much truncated ORF running from positions 266 to 1072. We have no evidence whether the short ORF in the West African isolates is expressed, but the absence of this ORF in the Zaire isolates indicates that it is not necessary for monkeypox virus transmission.

The base similarity of more than 99% in a gene sequence that is presumed to be redundant, implies that the separation of the various lineages has occurred fairly recently. It is the more remarkable then that so much variation was found in the form of insertions and deletions. We suggest that insertions and deletions occur at least as often as base substitutions during poxvirus DNA replication. Insertions and deletions, however, are more likely than base substitutions to disrupt the integrity of the gene product, and the resultant changes may be maintained only in sequences coding for gene products which are not required for continued transmission in the natural host.

The deletions shown in the monkeypox virus sequence were deduced by comparison to the corresponding sequences in variola and vaccinia viruses, but there is no evidence to suggest that either of these viruses is in the direct ancestral lineage of monkeypox virus.

We gratefully acknowledge the financial support of the Poliomyelitis Research Foundation. N. D. was employed by the SAIMR and M. R. was seconded from the State Vaccine Institute, Pinelands during the course of this project.
References


(Received 20 September 1993; Accepted 7 January 1994)