Independent Evolution of Monkeypox and Variola Viruses

NICOLA DOUGLASS AND KEITH DUMBELL*

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

Received 26 June 1992/Accepted 1 September 1992


We have now identified a sequence in monkeypox virus DNA which is a homolog of a 1,065-bp open reading frame in the conserved region of the variola virus genome but which has multiple deletions. This is strong evidence that monkeypox virus is not ancestral to variola virus and strengthens confidence in the long-term success of smallpox eradication.

Published restriction maps (1, 3, 4, 8) show that the majority of the HindIII cleavage sites in the central parts of the genomes are conserved between monkeypox virus and variola virus, and the respective maps can be aligned on this basis. Within the coding region, the variola virus Harvey HindIII D fragment is somewhat longer than the corresponding monkeypox virus Denmark HindIII E fragment (1, 8). Using variola virus Harvey HindIII D cloned into pAT153 and monkeypox virus Denmark HindIII E cloned into pBR329, we showed by hybridization and DNA sequencing that each end of variola virus HindIII D was homologous to the corresponding end of monkeypox virus HindIII E and was unrelated to the HindIII fragments on either side of monkeypox virus HindIII E. The location, within variola virus HindIII D, of the inferred extra sequence, not present in monkeypox virus, was first sought by a comparison of restriction site maps of the two HindIII fragments (Fig. 1). The locations of EcoRI and BamHI sites in variola virus HindIII D were known from unpublished work by Hamilton et al. (6a); these sites are compared with the locations of the corresponding sites for monkeypox virus HindIII E in Fig. 1a. The major discrepancy was between the two Eco-Bam fragments 2.9 and 2.4 kb long, respectively. These fragments were subcloned into pUC19 for more detailed study. Cleavage sites for the endonucleases KpnI, NcoI, and HindII were located in equivalent positions in the left half of each fragment. An Rsal site near the center of the variola virus fragment was not present in monkeypox virus. Further comparison of the right halves of the fragments was made by heteroduplex analysis. The two recombinant DNAs were linearized by digestion at the single KpnI site, denatured, mixed, allowed to reanneal, and treated with S1 nuclease for 4 min. The resulting fragments were separated on a 0.8% agarose gel at 2 V/cm. The regions 1.4 kb long at the right end of each insert were resistant to S1 digestion and were taken to be almost identical (data not shown). These results, summarized in Fig. 1b, appeared to localize the surplus variola virus sequence to a region around the Rsal site. The length of the HindII-Rsal fragment of the variola virus sequence in this region was reduced by exonuclease from the HindII end to 81 bp (as determined by subsequent sequencing) and cloned into pUC19 to be used as a hybridization probe. This probe failed to hybridize to the monkeypox virus HindIII E fragment or to any other part of the monkeypox virus genome. Further hybridization tests confirmed that this 81-bp sequence was not to be found in genomic DNA from each of six isolates of monkeypox virus (two each from West Africa, Nigeria, and Zaïre). In contrast, the probe hybridized to DNA from five strains of vaccinia virus, to DNA from one strain each of cowpox, camelpox, and taterapox viruses, and to cloned DNA from a second strain of variola virus (Somalia).

The 2.9-kb EcoRI-BamHI fragment of variola virus was reduced by about 1.4 kb from the Bam end by exonuclease and recloned for sequencing. Sequencing was done in both directions by using the dideoxy method of Sanger, modified for use with Sequenase enzyme and with double-stranded templates and synthetic oligonucleotides used as forward and reverse primers. Further subclones of reduced length were prepared by exonuclease digestion and recloning, as required by the initial sequencing results.

Figure 2 shows the variola sequence read to the right from the NcoI site (Fig. 1b). The probe sequence was found to be part of a 1,065-bp open reading frame (ORF) in the cloned DNA of variola virus Harvey, beginning at base 143 and running rightwards to base 1207 (Fig. 2). An appropriate subclone of the HindIII D clone of a second strain of variola virus (Somalia) was sequenced and found to contain an ORF identical to that of Harvey with the exception of a single nucleotide: the substitution of A for G at base 337 (Fig. 2).
Harvey is a strain of variola major virus, and Somalia is a strain of African variola minor virus (2).

Similar methods were used to obtain a corresponding sequence from monkeypox virus Denmark DNA. This could be aligned to the variola virus ORF with 96% base similarity by introducing four deletions of 24, 1, 391, and 24 bp (Fig. 2). The first of these deletions occurred 16 bases after the initial ATG in the variola virus sequence and was followed 40 bases later by a G-to-T substitution which introduced a stop codon into the monkeypox virus sequence. Twelve bases later is another in-frame ATG, from which the monkeypox virus sequence is similar to the variola virus sequence for 322 bases up to the single-base frameshift deletion at base 600 (Fig. 2). The next 68 bases retain a close similarity to variola virus up to the beginning of the major, 391 bp deletion, which is closely followed by two in-frame stop codons and 14 bases later is followed by the second 24 bp deletion. It was noted that the variola virus sequence corresponding to the 24 bp deletions was flanked by short, direct repeats, of which a single copy was present in the monkeypox virus sequence. This also implied that each of these deletions could be either placed as shown or shifted 6 bases along.

No sequencing of DNA from other strains of monkeypox virus was done, but the deletions shown by the hybridization experiments were confirmed for three other strains by using the polymerase chain reaction to amplify fragments spanning the region of the major deletion; in each strain, the resulting fragment was significantly shorter than the control fragment from variola virus. This experiment, done on genomic monkeypox virus DNA, also confirmed that there was no complete copy of this ORF elsewhere in the monkeypox virus genome.

The overall base similarity of 96% shows that there must be common ancestry of the DNA sequences examined in both variola and monkeypox viruses. In both major and minor forms of smallpox virus, this sequence had a fully conserved, presumptive translation product of 341 amino acids (40.5 kDa). In monkeypox virus, 60% of the presumptive product at the 3' end had been deleted or disrupted, and
there was also a truncation at the 5′ end, leaving a potential coding sequence less than one-third the length of that in variola virus.

It follows that variola virus and monkeypox virus have evolved independently since branching from their common ancestor and that variola virus could not be a direct descendant of the existing monkeypox viruses.

Comparison of the variola virus ORF with published vaccinia virus sequences shows that it is closely similar to the E5R ORF found in the Copenhagen strain of vaccinia virus (5, 6) and the equivalent ORF in vaccinia virus WR (GenBank accession no. M36339 and M59239). We sequenced the region equivalent to ESR in another strain of vaccinia virus (Dairen) and found a similar ORF.

It was noted that our variola virus sequence includes the equivalent of the beginning of the E4L ORF and the short E ORF B found in vaccinia viruses (5, 6). The latter lies wholly within the 391-bp sequence which forms the main deletion found in monkeypox virus DNA.

No function has currently been assigned to vaccinia virus ESR (5, 6). The presumptive translation products of the two variola virus ORFs differed at 14 residues from the corresponding residue in each of the three vaccinia virus ORFs. These 14 residues were conserved through the three vaccinia virus strains, but the vaccinia viruses varied among themselves at 10 other residues. At each of these 10 sites, the variola virus residue matched the residue in at least one of the vaccinia virus strains. A fuller report on the phylogeny of the ESR gene is in preparation.

Although the ESR sequence lies within the conserved central part of the Orthopoxvirus genome, its disruption in monkeypox virus suggests that it is a nonessential gene. This possibility is being tested with vaccinia virus.

**Nucleotide sequence accession numbers.** GenBank accession numbers and virus designations for the DNA sequences reported here are as follows: M95532, Harvey, Orthopoxviridae, variola virus, 1,560 bp; M95533, Somalia, Orthopoxviridae, variola virus, 1,560 bp; M95534, MPDENMAR, Orthopoxviridae, monkeypox virus, 1,148 bp; and M95535, VVDIE, Orthopoxviridae, vaccinia virus, 1,579 bp.

We thank A. Hamilton and P. J. Greenaway for the HindIII clones of variola virus DNA and for mapping EcoRI and BamHI sites on variola virus HindIII D.

We also thank the Poliomyelitis Research Foundation for financial support.

**REFERENCES**


