Proc. Natl. Acad. Sci. USA
Vol. 78, No. 3, pp. 1614–1618, March 1981
Biochemistry

Instability and reiteration of DNA sequences within the vaccinia virus genome

(tandem repeats/repetitive DNA/recombination/poxviruses)

BERNARD MOSS, ELAINE WINTERS, AND NORMAN COOPER

Laboratory of Biology of Viruses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205

Communicated by Wallace P. Rowe, December 15, 1980

ABSTRACT **The sequence arrangement within the nontranscribed portion of the inverted terminal repetition of the vaccinia virus genome exists in quasi-stable and unstable forms that are not distinguishable on the basis of viral infectivity. The unstable forms, which composed about 20% of a serially passaged stock of virus, were recognized by terminal heterogeneity on restriction endonuclease analysis. Instead of a single terminal fragment from each end of the genome, an array of eight or more fragments differing in size by 1650-base-pair increments was detected. This feature was not eliminated by repeated plaque purification, indicating that the population of DNA molecules with various numbers of reiterations can rapidly evolve from the DNA of a single virus particle. However, at each successive round of plaque purification, about 20% of the unstable isolates revert back to the more stable form. Stable forms are characterized by the presence of a set of 13–17 tandem 70-base-pair repeats on each side of a 435-base-pair intervening sequence near both ends of the genome. In contrast, the unstable forms possess sets of tandem repeats and intervening sequences that alternate many times in series. The transition between the two genomic forms and the evolution of the unstable form appear to be mediated by recombinational events.**

Nontranscribed, highly reiterated DNA arranged in tandem arrays of short or long repeats with internal periodicities is common in eukaryotic genomes (see ref. 1 for references). To highlight the possibility that the repeated DNA sequences may have no special function, the term "selfish DNA" was proposed (2, 3). The origin, perpetuation, and continued evolution of repeated DNA have been ascribed to unequal recombinational events (4). Evidence for reiteration of a nontranscribed segment of vaccinia virus DNA containing tandem repeats by such a process is presented here.

Vaccinia virus, the prototype of the poxvirus family, contains a 120-million-dalton or 180-kilobase-pair (kbp), linear, double-stranded DNA genome (5) with a 10-kbp inverted terminal repetition* (6, 7) that is divided into nontranscribed and transcribed portions (8, 9). The terminal 3-kbp nontranscribed region is of special interest because it contains a set of 13–17 direct tandem repeats of a 70-base-pair (bp) sequence on each side of a 435-bp intervening region (10). We suspected that this transposon-like (11) arrangement might promote recombinational events leading to terminal DNA sequence heterogeneity. Such a phenomenon could partly explain earlier reports (12, 13) of diffuse or submolar terminal restriction endonuclease fragments obtained upon analysis of serially passaged stocks of vaccinia virus.

During an analysis of individual vaccinia virus plaque isolates, several were found to have highly unstable genome structures. Multiple-length terminal fragments were obtained by restriction endonuclease analysis even after the virus was repeatedly plaque purified. Analysis of the DNA suggests that the bizarre unstable forms arise by reiteration of a DNA segment and occasionally revert back to more stable forms by unequal recombinational events that are a natural consequence of the genome structure of vaccinia virus.

MATERIALS AND METHODS

Virus. A stock of vaccinia virus (strain WR) originally obtained from the American Type Culture Collection and serially passaged in HeLa human cells at a multiplicity of one plaque-forming unit per cell was used. Plaque isolation was carried out on BSC-1 monkey cell monolayers, using a 1% agar overlay. Sufficient virus for DNA analysis was obtained by using one-fourth of the virus from a single plaque to inoculate BSC-1 monolayers in three to four 150-cm^2 bottles. Virus was purified by sucrose gradient sedimentation as described (14), except for the substitution of a Beckman SW 41 rotor. In some experiments, virus partially purified by sedimentation through a cushion of 36% sucrose was used.

Analysis of DNA. DNA was isolated from purified virus, digested with restriction enzymes, and analyzed by agarose gel electrophoresis as described (7, 10). Transfer of separated DNA fragments by blotting to nitrocellulose membranes (15) or preparative isolation of DNA by binding to glass powder (16) was according to established procedures. DNA labeled with ^{32}P by nick-translation (17) was hybridized to the immobilized DNA fragments as described (9).

Materials. Restriction endonucleases were purchased from Bethesda Research Laboratories and New England BioLabs. Radioisotopes came from Amersham.

RESULTS

Unstable Forms of Vaccinia Virus. To explore the basis of terminal heterogeneity in the vaccinia virus genome, individual plaques were picked from a serially passaged virus stock. These isolates were then used to infect BSC-1 monolayers at low multiplicity, and virus was purified to obtain sufficient DNA for restriction endonuclease analysis. Restriction fragments from the entire genome were separated by agarose gel electrophoresis and transferred to a nitrocellulose membrane. End fragments were identified by hybridization to a ^{32}P-labeled *Hpa* II DNA fragment containing the terminal 3.5 kbp of the genome. Control experiments were carried out, using our standard plaque-purified virus that had previously been picked from the stock described above. With the restriction enzyme *Xho* I, which cuts within the 10-kbp inverted terminal repetition, only the double molar 5.9-kbp G fragment derived from both ends of the genome was detected by hybridization to the probe (Fig.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); kbp, kilobase pair(s).

* We use the words "repetition" for the 10-kbp inverted terminal repetition, "repeat" for the tandem 70-bp repeating unit, and "reiteration" for the segment of DNA that is repeated in unstable virus isolates.

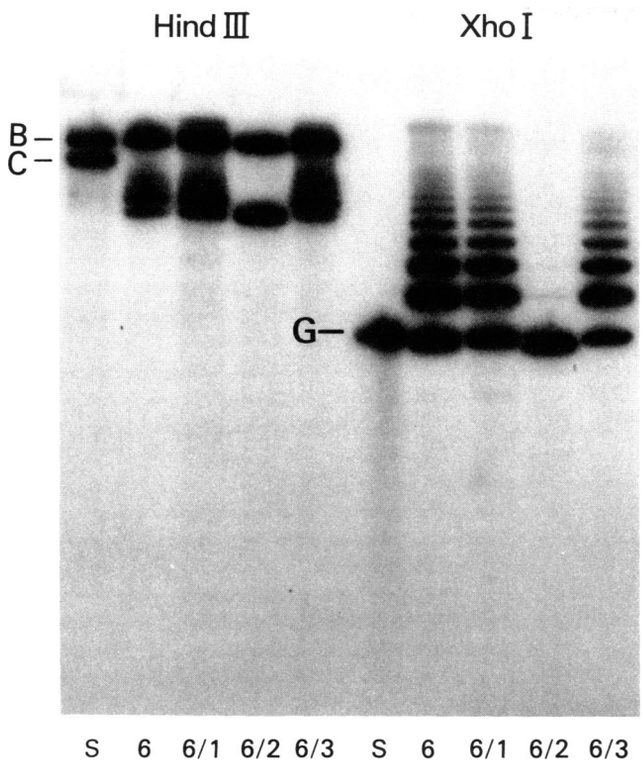

FIG. 1. Identification of terminal *Hin*dIII and *Xho* I restriction endonuclease fragments. *Hin*dIII and *Xho* I complete digests of DNA from the progeny of vaccinia virus plaque isolates was fractionated by 0.6% agarose gel electrophoresis, transferred to a nitrocellulose membrane, and hybridized to a ^{32}P-labeled DNA fragment probe. The latter was prepared by nick translation of a 3.5-kbp *Hpa* II fragment isolated from the end of the cloned 9-kbp *Eco*RI terminal genome fragment (8). After unhybridized DNA had been washed away, an autoradiograph was prepared. S, standard plaque-purified vaccinia virus used in this laboratory; 6, a randomly picked plaque isolate; 6/1, 6/2, and 6/3 are second-generation plaques obtained by replaquing 6.

1, track S). With restriction enzyme *Hin*dIII, which does not cleave within the repetition, the two large terminal B and C fragments of approximately 27 kbp and 20 kbp were detected by the hybridization procedure (Fig. 1, track S).

When 20 additional virus isolates were screened in this manner, many exhibited small deviations in the size of the terminal 5.9-kbp *Xho* I G fragment, which could be due to variations in the precise number of 70-bp repeats. Remarkably, four of the isolates examined appeared to have a series of various sizes of terminal *Xho* I fragments instead of one. This report is devoted to the analysis of one such bizarre variant.

The terminal fragments of a complete *Xho* I digest of plaque isolate 6 DNA appeared as an array or ladder in which the lowest rung was similar in size to the single 5.9-kbp terminal fragment obtained with standard virus (Fig. 1, track 6). The next higher band was the most intensely labeled one and was approximately 7.5 kbp. Each band above that one was slightly less intense and was also 1.6–1.7 kbp larger. In total, at least eight bands were counted. Densitometry indicated that the "normal" 5.9-kbp size terminal *Xho* I fragment composed 27% of the total and the successively higher bands composed 34, 22, 8, 5, 2, and less than 1%, respectively.

Because the two strands of vaccinia virus DNA are covalently linked at their ends (5), terminal restriction fragments have the unique property of rapid renaturation (5, 18). To determine whether the multiple terminal fragments of DNA from isolate 6 contain covalently linked ends, *Xho* I digests were denatured with alkali, neutralized, and treated with the single-strand nuclease S1. Analysis of the products by agarose gel electrophoresis indicated that at least the three lowest terminal fragments of the ladder had rapidly renatured and were resistant to S1 nuclease. Presumably, all of the terminal fragments had this property; however, the ethidium bromide stain was too faint to detect higher molecular weight bands. By contrast, only the 5.9-kbp fragment of standard virus DNA was resistant to nuclease S1 under these conditions.

Examination of a complete *Hin*dIII digest revealed that this plaque isolate also had a 9-kbp deletion that shortened the terminal *Hin*dIII C fragment. Elsewhere, we will demonstrate that this deletion, which does not appear to be directly related to the terminal heterogeneity of present interest, starts about 10.3 kbp from the left end of the genome. Significantly, however, a series of bands was also detected above the shortened left terminal *Hin*dIII C fragment (Fig. 1, track 6). Similarly, the blur above the full-length right terminal *Hin*dIII B fragment could be resolved into a series of bands when the electrophoresis was carried out for prolonged times (not shown).

The bizarre characteristics of this variant were most readily explained by the existence of a family of DNA molecules with a variable number of reiterations of a 1.65-kbp sequence at both ends. However, because the isolate had been plaque purified and passaged only once, the virus must be highly unstable. To check this, some of the remaining virus from the initial plaque number 6 was dispersed by sonication, diluted, and replaqued. The restriction endonuclease analysis of the DNA from the progeny of three daughter plaques is also shown in Fig. 1. With isolates 6/1 and 6/3, an array of terminal fragments virtually identical to those of 6 was detected, whereas 6/2 had a single labeled *Xho* I band of approximately 5.9 kbp.

This type of analysis was extended as indicated in Fig. 2. The isolates outlined by boxes were similar to 6/2 and had a single size terminal *Xho* I fragment. The other isolates, however, gave a ladder pattern virtually identical to that of 6/1. Because the ladder pattern was still obtained with virus that had been subjected to four successive plaque purifications, we concluded that the entire family of variable-length DNA molecules could be rapidly generated from the DNA in a single particle. The more stable forms that continuously occur with a frequency of

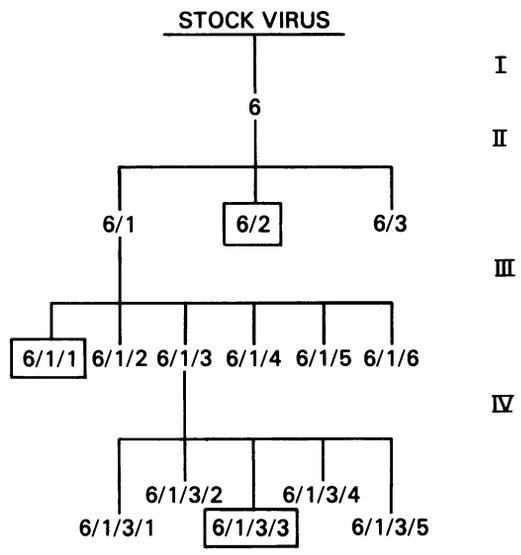

FIG. 2. Pedigree of plaque isolates. Isolates that are boxed have "stable" genotypes indicated by a single 5.9-kbp *Xho* I terminal DNA fragment. All other isolates have multiple terminal DNA fragments differing in length by 1.65 kbp.

about one in five are apparently revertants. Further analysis demonstrated that 6/2 and all the other stable revertants still have shortened *Hin*dIII C fragments. Therefore, the deletion on the left side of the genome does not bear a direct relationship to the phenomenon of instability. To further demonstrate the relative stability of the revertants, 10 daughter plaques were obtained from 6/1/3/3, and DNA from their progeny was analyzed: each had a single 5.9-kbp *Xho* I terminal fragment.

Identification of the Reiterated Sequence. From the site of *Xho* I cleavage we concluded that reiterations of the 1.65-kbp sequence occur within the terminal 5.9 kbp of the vaccinia virus genome. Because the result obtained with *Sal* I, which cuts 3.5 kbp from the end of the vaccinia virus genome, was similar to that obtained with *Xho* I, further localization of the reiterated DNA was achieved (not shown).

A restriction endonuclease map of the terminal 4 kbp of the vaccinia virus genome showing *Sal* I, *Dde* I, and *Alu* I sites (10) is presented in Fig. 3*B*. Each of the two enlarged areas enclose a set of 13–17 direct tandem repeats of a 70-bp sequence containing an *Hin*fI restriction site. The last *Hin*fI site of the first set of repeats and the first of the second set are separated by a 435-bp intervening sequence containing single *Dde* I and *Alu* I sites.

When our standard virus DNA was digested with *Dde* I or *Alu* I, a large number of restriction fragments were resolved by agarose gel electrophoresis. However, when those fragments were transferred to a nitrocellulose membrane and probed with purified ^{32}P-labeled 70-bp *Hin*fI fragment containing the repeat sequence, only two were labeled. As shown in Fig. 3*A* (track S), the two *Alu* I bands were approximately 1340 and 2560 bp and the two *Dde* I bands were 1140 and 2825 bp. With other plaque isolates, we found differences in the sizes of these bands consistent with variations in the precise number of tandem repeats of the 70-bp sequence. The map in Fig. 3*B* indicates that the smaller *Dde* I and *Alu* I fragments were derived from the very end of the genome, whereas the larger ones were derived from the adjacent sequence.

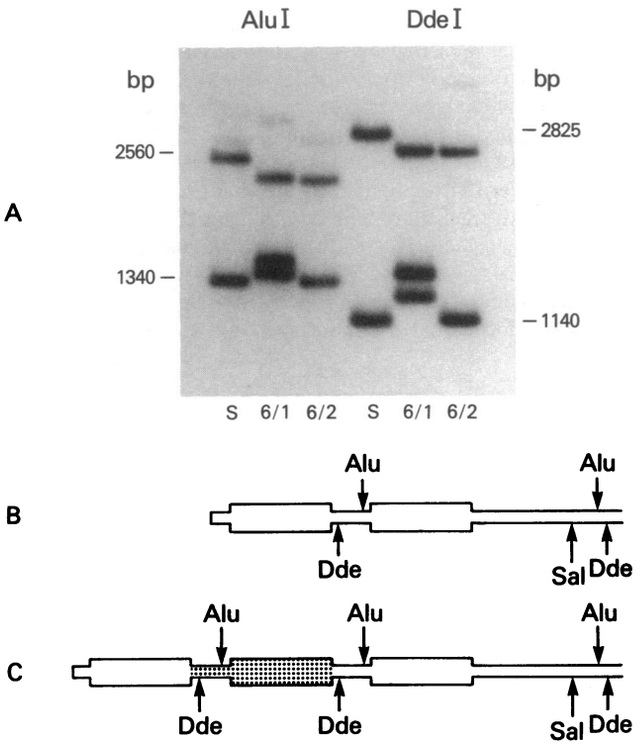

FIG. 3. Identification of the reiterated sequence. (*A*) DNA of standard virus (S) or plaque isolates (6/1, 6/2) were digested with either *Dde* I or *Alu* I, electrophoresed on a 1% agarose gel, transferred to a nitrocellulose membrane, and hybridized to a purified 70-bp repeat fragment labeled with ^{32}P. The latter was obtained from an *Hin*fI digest of cloned DNA. An autoradiograph is shown. (*B*) Restriction endonuclease map of terminal 4 kbp of the vaccinia virus genome (10). Each set of 13–17 tandem 70-bp repeats is indicated by the enlarged segments. (*C*) Restriction endonuclease map of terminal fragments containing reiterated DNA. The reiterated DNA is indicated by the segment filled with dots.

When DNA from 6/1 was examined in the same manner, three bands were detected with both *Alu* I and *Dde* I (Fig. 3*A*). The absence of a ladder pattern indicated that both restriction enzymes cleaved within the reiterated unit. With *Alu* I, the three bands were estimated to be 2300, 1525, and 1425 bp; and with *Dde* I, they were 2575, 1425, and 1275 bp. That none of these bands were identical in size to the ones from standard virus was consistent with the variations in number of 70-bp repeats alluded to above. For example, the upper *Alu* I and *Dde* I bands of 6/1 are both about 260 bp smaller than the respective bands from our standard virus, suggesting that the second set of tandem repeats in 6/1 contains 13 rather than 17 units. The more significant finding, however, was the presence of a third band, which must be derived from reiterated DNA. Furthermore, the reiterated DNA must contain both *Dde* I and *Alu* I sites in order to generate a third fragment with each enzyme, and it must contain 70-bp repeats for the fragment to hybridize to the probe.

In the model shown in Fig. 3*C*, the reiterated DNA is indicated by the dotted area and is composed of a copy of the intervening sequence and one set of repeats. Assuming that set of tandem repeats contains 17 units of the 70-bp sequence and the intervening sequence is 435 bp, the size of the reiterated DNA would be about 1.63 kbp, in excellent agreement with the increment in size of the terminal *Xho* I fragments. In DNA molecules with multiple reiterations, the entire dotted area in Fig. 3*C* would be tandemly repeated. Although this would not generate different size *Dde* I or *Alu* I fragments, the extra band would be multimolar. Indeed, densitometry of the autoradiograph showed that the 1425-bp *Dde* I band and the 1425-bp *Alu* I band were most intense, indicating that they are derived from the reiterated DNA. Furthermore, the identity in sizes of the *Dde* I and *Alu* I bands containing the reiterated DNA is consistent with the model in Fig. 3*C*. The sequence arrangement in Fig. 3*C* was also demonstrated by electrophoresis of the partial *Dde* I digestion products of the 7.5-kbp terminal *Xho* I fragment of 6/1 DNA labeled with ^{32}P at the latter restriction site (not shown).

As a further test of this model, we probed *Dde* I and *Alu* I digests of revertant DNA to see if they had lost the 1425-bp band. The result obtained with 6/2 DNA is shown in Fig. 3*A*. Significantly, the 1425-bp band was absent. We were surprised, however, to note that the terminal *Dde* I and *Alu* I fragments of 6/2 were the same size as in our standard virus and different from 6/1. Because of the possibility that the size change was associated with reversion, we analyzed the other revertants 6/1/1 and 6/1/3/3, which, unlike 6/2, are directly descended from 6/1. Again, the 1425-bp *Dde* I fragment was missing, but now the remaining two fragments were identical in size to the 1275- and 2575-bp fragments obtained from 6/1 (Fig. 4). This important result supported our belief that the revertants are direct descendants of the unstable form and that the events that lead to reversion do not specifically involve changes in the number of tandem repeats of the 70-bp sequence or involve other measurable insertions or deletions in the sequence.

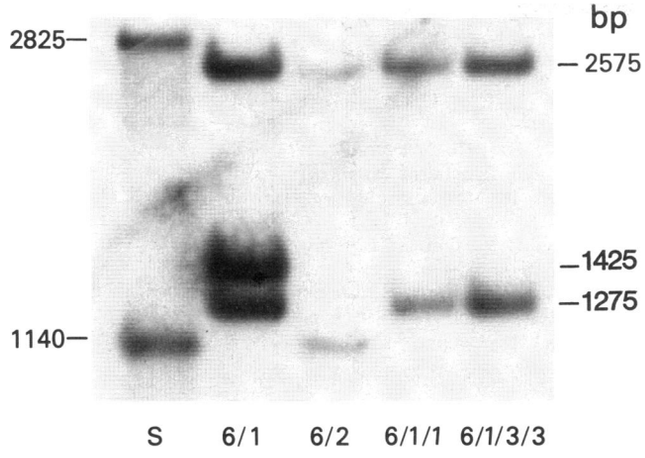

FIG. 4. Hybridization of ^{32}P-labeled 70-bp repeat to *Dde* I fragments of vaccinia DNA. Hybridization was carried out as in Fig. 3*A*. Pedigree of isolates is shown in Fig. 2.

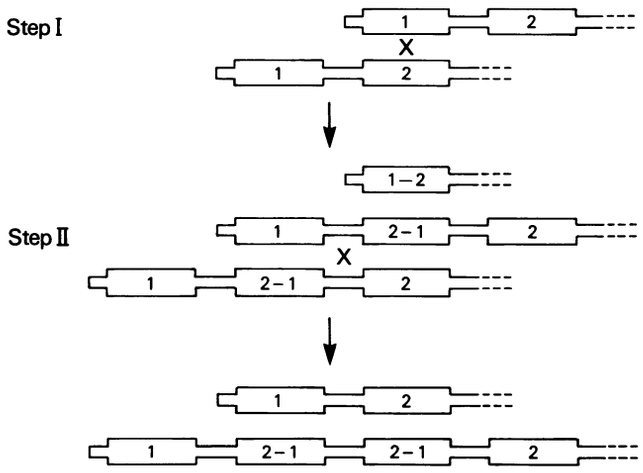

FIG. 5. Recombinational model for the generation of unstable variants and reversion to more stable forms by unequal crossover. Ends of DNA molecules are represented. Each numbered segment encloses a set of 13–17 direct tandem repeats. In step I, recombination occurs between two sets of repeats to generate the unstable variant. In step II, recombination occurs within intervening regions (as shown) or between sets of tandem repeats. Further recombinational events lead to additional reiterations as well as to the original genome structure.

DISCUSSION

Within the vaccinia virus genome there is a 10-kbp inverted terminal repetition (6, 7) containing two sets of 13-17 tandem repeats of a 70-bp sequence separated by an intervening sequence of approximately 435 bp (10). The present study indicates that, in about 20% of the plaque isolates from a stock of vaccinia virus, one set of short tandem repeats and the intervening sequence form a larger 1.65-kbp tandemly repeated sequence. Extreme instability is a novel characteristic of these variants; even after repeated plaque purification, restriction endonuclease analysis of the DNA reveals an array of terminal fragments with 1.65-kbp incremental differences in length due to variations in the number of copies of the reiterated sequence. Thus, a population of DNA molecules with various numbers of reiterations can rapidly evolve from the DNA of a single virus particle. In addition, revertants with the stable predominant genome structure arise continuously. These revertants are not contaminants from the original virus stock, because they occur with the same frequency after repeated plaque purifications of the unstable variants. Moreover, the revertants and unstable variants have the same 9-kbp deletion, consistent with a common ancestry.

A recombinational model accounting for the generation and properties of the highly unstable variants and their reversion to the more stable predominant form is presented in Fig. 5. The first step is unequal recombination between the first set of tandem repeats of one DNA molecule and the second set of another. (Because this sequence is within the inverted terminal repetition, recombination between the two ends of a single molecule may also be possible.) As shown, one of the recombinants formed in this step is a molecule that has deleted one set of tandem repeats and the intervening sequence. However, virus containing a genome of this type has not yet been isolated. The other recombinant contains three sets of tandem repeats and a duplication of the previously unique intervening region. Now, recombination could occur between two DNA molecules of the latter type or between the latter and the original types. Moreover, the rate of unequal crossover events would be further enhanced by the juxtaposition of multiple short and long tandem repeats. If two DNA molecules were to align in the mismatched fashion shown in step II, there would be a 2.5-kbp stretch of perfect homology. Crossover between the duplicated intervening regions could generate DNA molecules with four sets of tandem repeats and a triplication of the intervening sequence. Because the reiterations occur within the inverted terminal repetition, recombination could occur between opposite ends of different molecules or even the same molecule. This process could go on to generate a large family of DNA molecules with many reiterations at both ends. However, as shown in Fig. 5, the same recombinational events also generate revertants with the original and more stable genome structure.

The above model accounts for the spontaneous generation of the extremely unstable forms, the accelerated rate of unequal crossover events within the latter, and the reversion to the original genotype. Moreover, the only process required is recombination, which from genetic studies with poxviruses is known to occur at a high rate (19). Although it may be possible to devise alternative models in which reiterations occur during DNA replication, too little is known regarding that process to consider them at this time. Other possibilities such as a mutation leading to accelerated rates of recombination cannot be ruled out. However, the frequency of revertants is far too high for back mutations.

End fragments containing at least seven reiterations of the 1.65-kbp sequence were detected. Although the majority of DNA molecules must have reiterations at both ends, we do not know whether the same number is present at the two ends of a single molecule. Clearly, this would be required if the symmetry of the genome were to be preserved (20). From the intensity of the autoradiographs, it was determined that ends containing a single reiteration were most abundant and that ends containing additional reiterations were present in successively lower amounts. However, only DNA packaged in virions was examined, and the possibility that a different distribution is present within the replicating pool has not been excluded. The greater than 9-kbp deletion on the left side of the genome of the unstable isolate examined could compensate for the added length caused by the reiterations if there are packaging constraints. Alternatively, the presence of both aberrations in a single isolate could be coincidental. In this regard, analysis of another independently isolated unstable virus isolate did not reveal a large deletion.

Remarkably, when plaque sizes, particle to plaque-forming unit ratios, and burst sizes of the unstable variant containing

the 9-kbp deletion, the stable revertant, and standard virus were compared, no major differences were detected. Elsewhere, we will show that the deleted region encodes at least eight early and three late polypeptides, which evidently are not required for efficient replication in HeLa cells. Large deletions of this type have also been found in rabbitpox (21), cowpox (22), and vaccinia virus.[†]

In considering the possible mechanisms involved in reiteration of the vaccinia virus DNA sequence, it would be of interest to know whether a similar phenomenon would occur if that segment of DNA were introduced into plasmids or eukaryotic genomes. If the arrangement of tandem repeats and intervening sequences promotes a high rate of unequal crossover events, as we have suggested, then reiteration might occur unhindered by any packaging constraints of the virus. Moreover, because these noncoding vaccinia virus DNA sequences are unlikely to provide any selective advantage to their hosts, their amplification would be consistent with the selfish DNA hypothesis (2–4).

We thank Ernest Barbosa for carrying out some preliminary experiments, Bahige Baroudy for samples of purified 70-bp repeat fragment, Riccardo Wittek for advice on experimental procedures, and Jeanne Carolan and members of the Editorial Office, National Institute of Allergy and Infectious Diseases, for assistance with the preparation of the manuscript.

[†] Panicalli, D. L. & Paoletti, E. (1980) in *Abstracts of the Annual Meeting of the American Society for Microbiology*, Miami Beach, FL, p. 228.

1. Donehower, L. & Gillespie, D. (1979) *J. Mol. Biol.* **134**, 805–834.
2. Doolittle, W. F. & Sapienza, C. (1980) *Nature (London)* **284**, 601–603.
3. Orgel, L. E. & Crick, F. H. C. (1980) *Nature (London)* **284**, 604–607.
4. Smith, G. (1976) *Science* **191**, 528–535.
5. Geshelin, P. & Berns, K. I. (1974) *J. Mol. Biol.* **88**, 785–796.
6. Wittek, R., Menna, A., Muller, H. K., Schumperli, D., Boseley, P. G. & Wyler, R. (1978) *J. Virol.* **28**, 171–181.
7. Garon, C. F., Barbosa, E. & Moss, B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4863–4867.
8. Wittek, R., Barbosa, E., Cooper, J. A., Garon, C. F., Chan, H. & Moss, B. (1980) *Nature (London)* **285**, 21–25.
9. Wittek, R., Cooper, J. A., Barbosa, E. & Moss, B. (1980) *Cell* **21**, 487–493.
10. Wittek, R. & Moss, B. (1980) *Cell* **21**, 277–284.
11. Calos, M. P. & Miller, J. H. (1980) *Cell* **20**, 579–595.
12. Wittek, R., Muller, H. K., Menna, A. & Wyler, R. (1978) *FEBS Lett.* **90**, 41–46.
13. McCarron, R. J., Cabrera, C. V., Esteban, M., McAllister, W. T. & Holowczak, J. A. (1978) *Virology* **86**, 88–101.
14. Boone, R. F. & Moss, B. (1978) *J. Virol.* **26**, 554–569.
15. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683–3687.
16. Vogelstein, B. & Gillespie, D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 615–619.
17. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
18. DeFilippes, F. M. (1976) *J. Virol.* **17**, 227–238.
19. Fenner, F. (1969) *Curr. Top. Microbiol.* **48**, 1–28.
20. McFadden, G. & Dales, S. (1979) *Cell* **18**, 101–108.
21. Moyer, R. W. & Rothe, C. T. (1980) *Virology* **102**, 119–132.
22. Archard, L. C. & Mackett, M. (1979) *J. Gen Virol.* **45**, 51–63.