JOURNAL OF VIROLOGY, Feb. 1984, p. 356–362
0022-538X/84/020356-07$02.00/0

Vol. 49, No. 2

Unstable Heterozygosity in a Diploid Region of Herpes Simplex Virus DNA

SUSANNAH L. VARMUZA* AND JAMES R. SMILEY
Pathology Department, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

Received 30 June 1983/Accepted 7 October 1983

We have examined the behavior of a herpes simplex virus strain KOS isolate in which the two inverted repeats flanking the short segment of viral DNA differ in length by approximately 60 base pairs. We find that individual viral DNA molecules exist which contain the two distinguishable repeats, demonstrating that heterology between the repeats is tolerated. However, viruses heterozygous for the two different repeats are unstable, segregating both classes of homozygotes at a high frequency. We propose that this segregation is a consequence of the high-frequency recombination events which also result in genome segment inversion.

The genome of herpes simplex virus (HSV), a linear duplex DNA of approximately 100×10^6 daltons, displays an unusual topography (Fig. 1) in which two unique DNA sequences, one long (UL) and one short (Us), are both flanked by fairly substantial inverted repeats designated b and c, respectively (4, 15, 16, 21). In addition, the genome is terminally redundant, and this terminal direct repeat (designated a) is also present in an inverted orientation at the L-S junction (Fig. 1). As originally proposed by Sheldrick and Berthelot (15), DNA extracted from virions consists of an equimolar mixture of four isomers, differing in the relative orientations of the long (L) and short (S) DNA segments (4, 16, 21). Although the mechanism of segment inversion remains unclear, it likely depends on frequent intramolecular recombination events between the inverted repeats present at the termini and L-S junction (3, 11, 12, 15, 17). Recent evidence suggests that the "a" sequence is both necessary and sufficient for these recombination events (2, 13).

The diploid inverted "c" sequences flanking Us contain the gene for ICP4 (1, 9, 20), an immediate-early polypeptide believed to be essential for the expression of delayed-early and late viral genes. Somewhat surprisingly, temperature-sensitive mutations affecting this diploid gene, which are necessarily present in both copies of the gene, are recovered at a reasonable frequency (14). As it is extremely improbable that both copies of the gene acquire the same mutation independently, this observation suggests that a mechanism exists which generates mutant homozygotes from wild-type/mutant heterozygotes at a high frequency. This suggestion is consistent with the results of intertypic marker rescue of ICP4 mutations of HSV-1 by HSV-2 DNA fragments: Knipe et al. (7) found that of 10 rescue products, 8 were heterodiploids and 2 were homozygous for the HSV-2 rescuing DNA sequence. As the two homozygous HSV-2 insertions mapped within the "c" sequence close to the a sequence, and the remaining heterozygous replacements mapped farther from the "a" sequence, the authors concluded that the "a" sequence, and the "c" sequences immediately adjacent to it, were "obligatorily identical." However, it is obvious that these observations do not demonstrate that these sequences are obligatorily identical: they demonstrate only that changes in one repeat can, at some frequency, appear in the other. Recently, Davison and Wilkie (3) have demonstrated that heterotypic "a" sequences are tolerated in HSV DNA. In this report, we show that heterologies are also tolerated in the "c" sequences immediately adjacent to the "a" sequence. However, we also find that viruses bearing such heterozygous "c" sequences segregate progeny of both classes. These results demonstrate directly that the "c" sequences of freely isomerizing HSV DNA are not obligatorily identical, but that a mechanism exists which generates homozygotes from heterozygotes at a high frequency. We suggest that this process depends on the high-frequency recombination events initiated within the "a" sequence.

* Corresponding author.

MATERIALS AND METHODS

Virus and cells. HSV-1 strain KOS and Vero cells were used throughout this study.

Plaque purification of virus. Two methods of plaque purification were employed in this study. The first involved an endpoint dilution of the virus stock. Confluent monolayers of Vero cells (2 cm^2) in 24-well plates (Linbro) were infected with 0.1 PFU per well. Infected cells were harvested 2 to 3 days later and frozen at −70°C. The second method involved the selection of thymidine kinase (TK)-deficient (TK^-) mutant clones of virus. Confluent monolayers in 100-mm petri dishes (Plastipak) were infected with 5×10^3 PFU of the virus stock. After 2 to 3 h of absorption, medium containing 100 μg of thymidine arabinoside (AraT, Calbiochem) per ml and 0.05% human immune serum was added to each plate. Two days later, the medium was removed, and individual plaques were lightly touched with a sterile wooden rod and transferred to 2-cm^2 monolayers of Vero cells. Infected cells were harvested after several days of incubation in the presence of 100 μg of AraT per ml and frozen at −70°C.

Labeling and extraction of viral DNA. Confluent monolayers of Vero cells in 25-cm^2 tissue culture flasks (Corning) were infected with 10^6 PFU of virus. After 2 to 3 h of absorption, 4 ml of phosphate-free minimal essential medium containing 100 μCi of ^{32}P-labeled orthophosphoric acid was added to each flask. One day later, infected cells were harvested, and the labeled viral DNA was selectively extracted by the method of Hirt (5). Occasionally, unlabeled viral DNA was prepared in a similar way. More often, unlabeled DNA was isolated from cytoplasmic nucleocapsids as previously described (18). After extraction, DNA was purified by phenol extraction and ethanol precipitation.

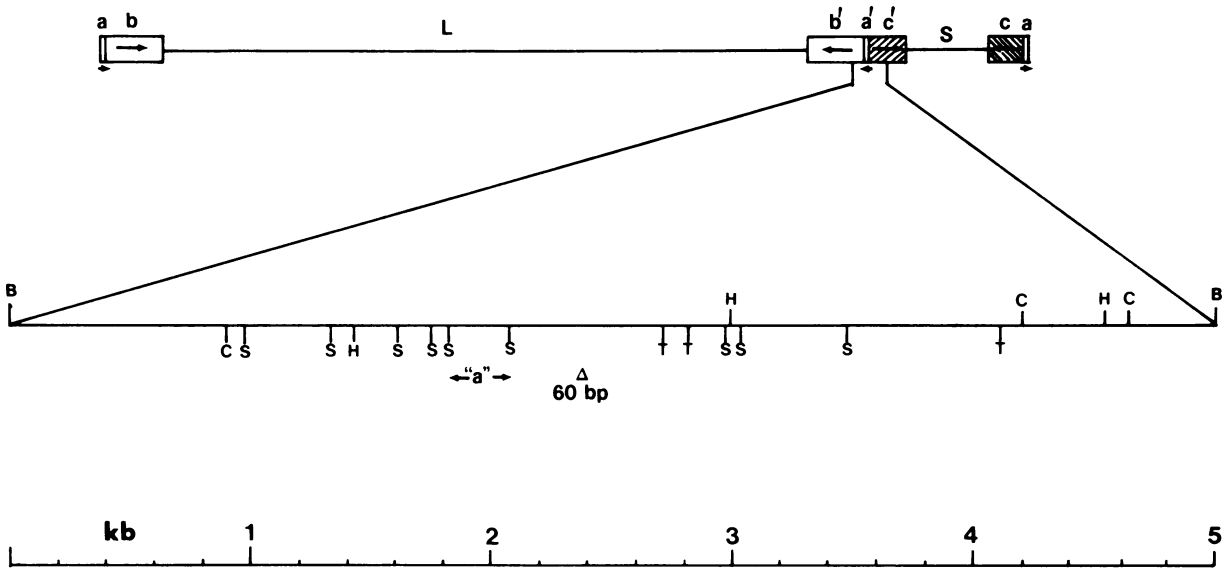

FIG. 1. Partial restriction map of HSV-1 (KOS) joint region. Some of the data were obtained from Wagner and Summers (19). The "a" sequence is putatively assigned to the *Sma*I fragment indicated, although the smaller *Sma*I fragment to the left may also be included. The deletion was mapped by Southern blot to the *Sma*I-*Taq*I fragment indicated. Precise localization awaits sequence analysis. Abbreviations: B, *Bam*HI; C, *Sac*I; H, *Hin*fI; S, *Sma*I; T, *Taq*I.

Restriction endonuclease cleavage, electrophoresis, and Southern blotting. Viral DNA was cleaved with various restriction endonucleases as indicated in the figure legends. Gel electrophoresis, Southern blotting (18), and hybridization were performed as previously described (17).

RESULTS

During the course of another study (19), a plaque-purified stock of HSV-1 strain KOS was uncovered which exhibited a heterozygosity in the joint-spanning and terminal fragments containing the diploid "c" sequence. The nature of the heterozygosity was an approximately 60-base pair difference in size which generated a doublet band in restriction digests. This stock, designated KOSC, was plaque purified by endpoint dilution to yield three progeny stocks: one contained only the smaller of the "c" containing fragments, one contained only the larger, and the third, KOSC1, contained the doublet.

An *Alu*I digest of the DNA from each of these stocks is shown in Fig. 2, illustrating several points. First, in all three virus stocks, the joint-spanning fragments migrate as a series of related bands, differing from one another by approximately 280 bases. These bands represent L-S junctions bearing increasing numbers of copies of the "a" sequence (19). In the case of C1, each of these successive bands is split, indicating that heterozygosity for the 60-base deletion occurs on all size classes of junctions. Second, as discussed in more detail below, the S terminal fragments reflect the structure of the L-S junction: when only the larger of the junctions is found (C2), only the larger terminus is present; when only the smaller junction is found (C3), only the smaller terminus is present; and when the junctions are heterozygous (C1), the termini are also heterozygous. Note that the larger terminus comigrates with another viral fragment, yielding a pattern reflecting the relative quantities of the termini: the intensity of the larger band is greater than that of the lower band where both termini are present (C1), whereas they are of equal intensity where only the lower form of the terminus is present (C3). The intensity of this band is greater than that of surrounding fragments in C2 where the terminus comigrates with the other band.

The presence of the doublet *Alu*I joint-spanning fragments of this stock was consistent with two possibilities: either the plaque purification had not resulted in clonal populations of virus, and the C1 stock arose from a double infection with two different homozygous viruses, or the C1 stock was in fact a clone, with heterozygous "C" sequences differing in length by 60 base pairs. To distinguish between these possibilities, KOS C1 was subjected to two further rounds of plaque purification, the first by endpoint dilution and the second by selecting TK-deficient mutant plaques.

Figure 3 shows a *Sac*I digest of labeled viral DNA from 14 different subclones of C1, isolated by endpoint dilution. Three of these 14 stocks were heterozygous, as evidenced by the presence of doublet joint-spanning fragments (indicated by arrows). These are shown in tracks marked 2-g, 2-k, and 2-r. All of the remaining subclones were homozygous for one or the other of the two forms. The S termini (indicated by arrows) were also heterozygous in these three subclones, although this is not as clearly demonstrated in *Sac*I digests because the smaller version of the S terminus comigrates with another unrelated *Sac*I fragment. In stocks homozygous for this form, all S termini comigrated with this other fragment, whereas in stocks homozygous for the larger form, all of the S termini were resolved from this band. As expected, the three heterozygous stocks displayed a one-half molar abundance of the larger form of the S terminus. These data suggest that 3 of 14 subclones of C1 retained the apparent parental heterozygosity. The possibility still remained, however, that these stocks were not clonally derived but were products of mixed infection.

To more rigorously ensure that presumed subclones were in fact derived from a single PFU, we performed another

round of plaque purification designed specifically to eliminate plaques arising from a double infection. This was done by selecting for TK^- mutant plaques, using AraT. The rationale for this approach was that TK^+ virus induces TK activity, which phosphorylates AraT, enabling its lethal incorporation into viral DNA. TK^- mutants of HSV are therefore resistant to AraT, because the drug is not phosphorylated. In a cell coinfected with TK^+ and TK^- viruses, the drug should be incorporated lethally into both TK^+ and TK^- DNA, preventing plaque formation. Since spontaneous TK^- mutants are relatively rare (10^{-3} to 10^{-4}, unpublished data), essentially every cell infected with only TK^- virus will necessarily be infected with only one infectious genome. Consequently, TK^- plaques are almost certainly clonal in origin (the probability of a mixed infection with two different TK^- viruses is ca. 10^{-6} to 10^{-8} of that of a TK^+/TK^- coinfection). The parental heterozygous stock used in this experiment was KOS C1r (see Fig. 3). From a total of 4×10^4 PFU plated out, 40 AraT-resistant, TK^- plaques arose, of which 31 were analyzed. Figure 3 shows a *Sac*I digest of labeled DNA from some of these TK^- isolates. Again, the heterozygosity persisted at a reasonable frequency (10 of 31 clones). Consequently, we are confident that these heterozygous stocks arose from single heterozygous DNA molecules. Therefore, we conclude that the KOS C1r stock, at the time it was subcloned, consisted of approximately one-third heterozygotes and two-thirds homozygotes.

Figure 4 shows *Bam*HI and *Alu*I digests of labeled DNA from representatives of all three generations of virus analyzed and Southern blots of similar digests probed with the *Sac*I joint-spanning fragment. The heterozygosity is more clearly displayed in the *Alu*I digests. Figure 5 depicts a "genealogical tree" of the viruses analyzed in these experiments.

The position of the 60-base deletion in the smaller of the two junctions was mapped to a 500-bp *Taq*I-*Sma*I fragment immediately adjacent to the "a" sequence (Fig. 6). The *Sac*I digests (on the right) indicate that the *Sac*I junction (dots) and S terminus (asterisks) contain the deletion. As the junction and S terminus are reduced in size by restriction enzymes that cut closer to the L-S junction, one can see that the heterozygosity is still present. For example, *Hin*fI junction fragments and S termini contain the deletion, as do *Sac*I-*Taq*I junction and S terminal fragments. *Sma*I digestion finally localizes the deletion to the 927-bp *Sma*I fragment immediately adjacent to the "a" sequence (see arrows). *Taq*I digestion of this fragment indicates that the deletion resides in the largest of the *Sma*I-*Taq*I fragments which maps adjacent to the "a" sequence (see arrowheads).

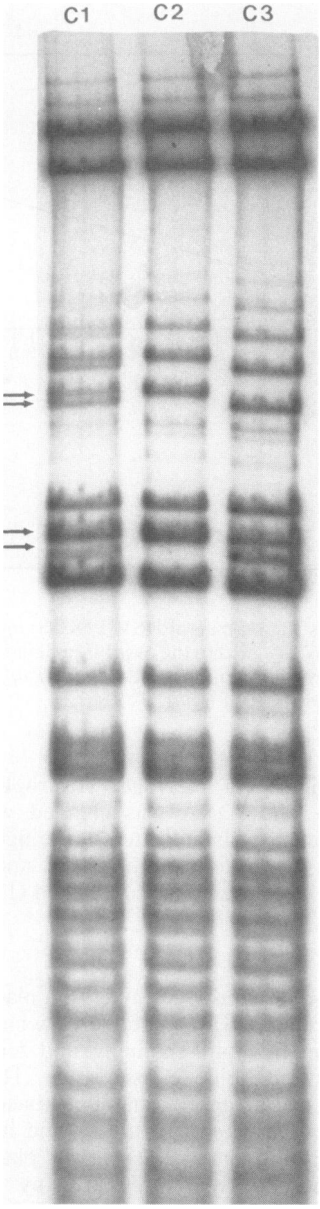

FIG. 2. *Alu*I digestion pattern of DNA from HSV-1 (KOSC) progeny. The heterozygote KOSC was plaque purified by endpoint dilution, and three progeny were expanded. After labeling the DNA in vivo, we compared the *Alu*I restriction patterns. Arrows indicate the two forms of the joint fragment (upper set) and S terminus (lower set).

DISCUSSION

There are two major conclusions that we have drawn from this work. First, the diploid "c" sequences immediately adjacent to the "a" sequence are not "obligatorily identical" (7). Our data demonstrate that individual HSV DNA molecules can bear two distinguishable "c" sequences. Since Davison and Wilkie (3) have shown that the "a" sequences are also not always identical, it appears that the previous conclusion of Knipe et al. (7), namely, that heterologies in the "a" and nearby "c" sequences are not tolerated, was mistaken.

Second, although sequence heterology between "c" sequences is tolerated, such heterozygotes are unstable. Virus stocks descended from a heterozygous virus particle contain approximately one-third heterozygotes after growth from 1 to 10^7 PFU. The remaining two-thirds of the virus are homozygous for either of the two classes of "c" sequence. From this observation, it follows that a mechanism exists which generates homozygous genomes from heterozygotes at a high frequency. The heterozygote can be viewed as an intermediate in a chain of events starting with the introduction of a change in one of the diploid sequences, followed by a later acquisition of this change in the other. This interpretation explains the observation of both heterodiploid and homodiploid recombinant forms of the ICP4 gene in intertypic marker rescue experiments (8), and the relative ease with which temperature-sensitive mutants of ICP4 are isolated.

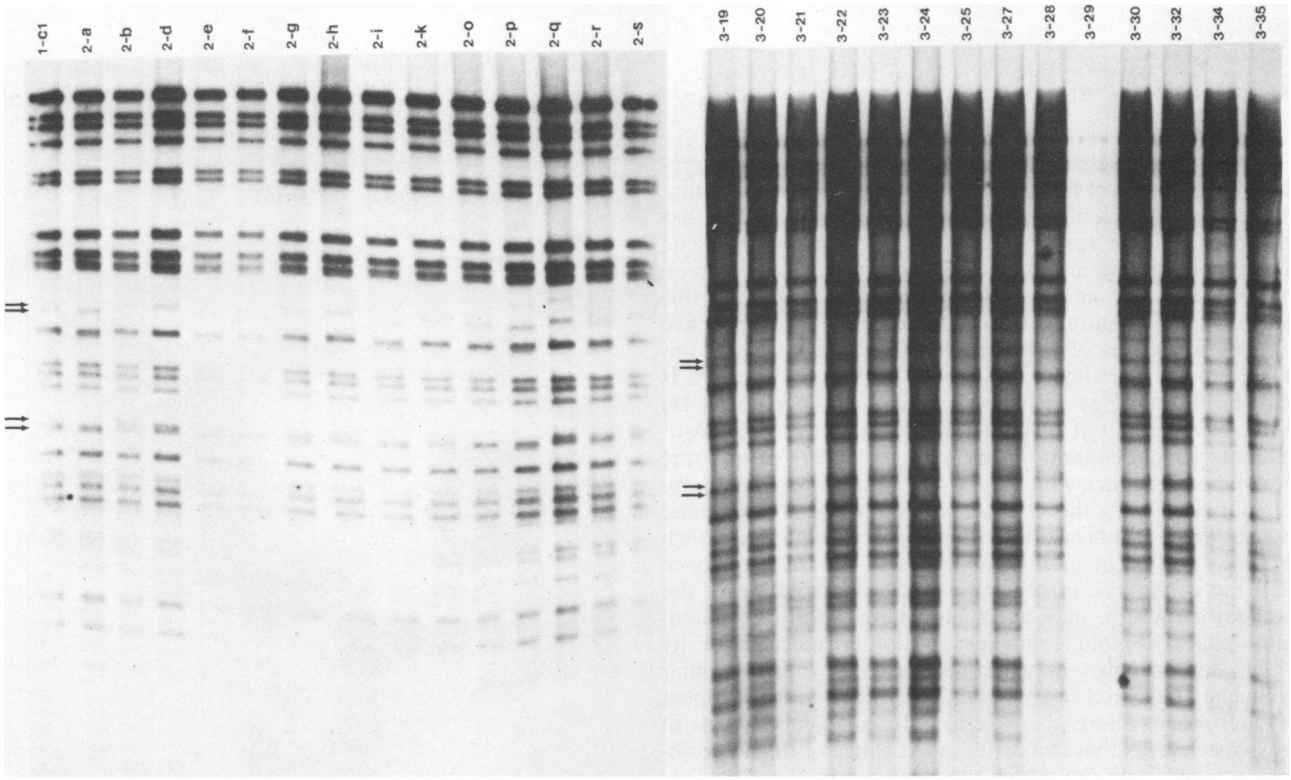

FIG. 3. *Sac*I digestion pattern of plaque-purified progeny. Heterozygous stocks were plaque purified by endpoint dilution (gel on left) or TK$^-$ selection (gel on right), and labeled DNA from progeny was examined as described in the legend to Fig. 2, except that *Sac*I was used instead of *Alu*I. Upper arrows indicate the two forms of the joint fragment; lower arrows indicate the two forms of the S terminus. The first number in the code above each track indicates the "generation," i.e., 1-c1 is first generation, 2-a is second generation, and 3-19 is third generation. KOSC1 (1-C1) was the parental stock from which all of the second-generation progeny were derived, and KOSC1r (2-r) was subjected to TK$^-$ selection to generate all of the third-generation progeny.

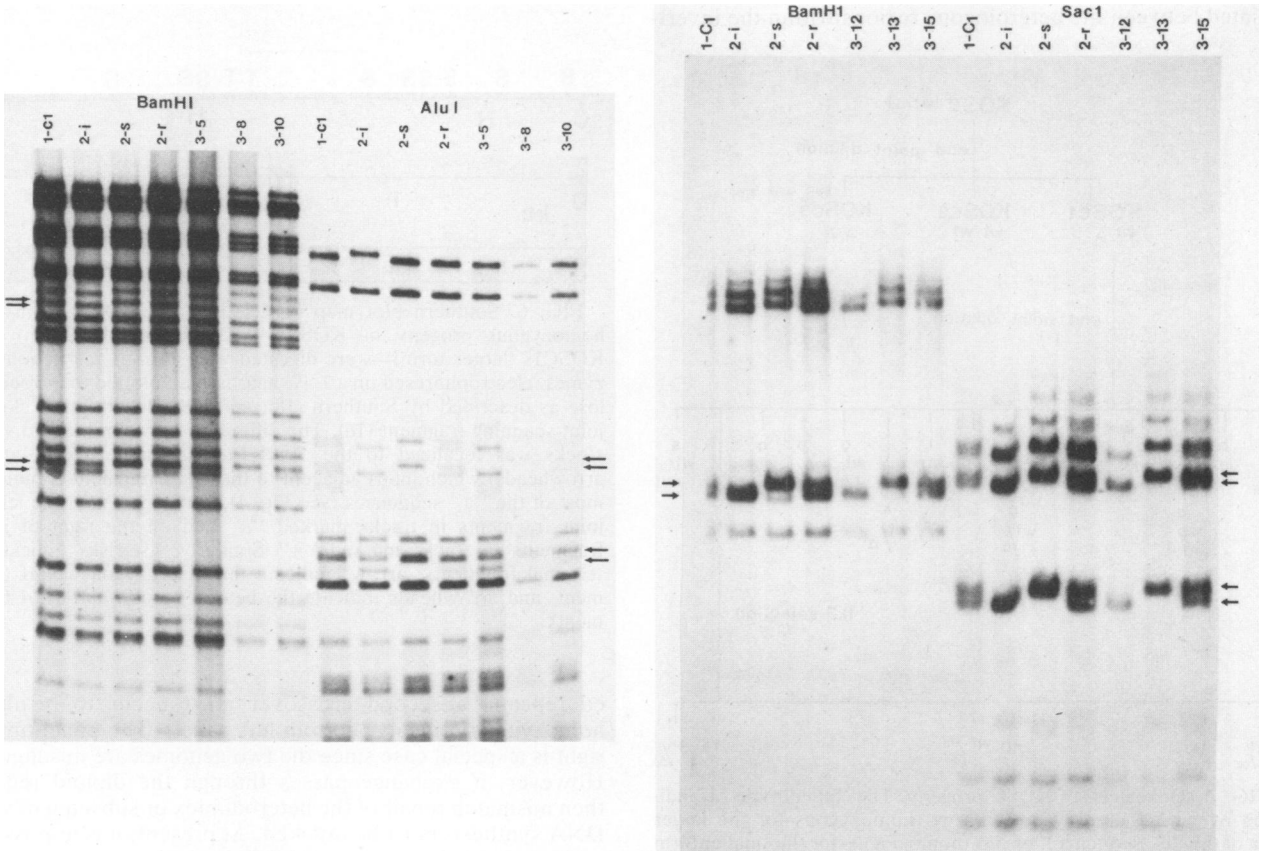

FIG. 4. Restriction profile of parental and plaque-purified progeny virus. DNA from representative stocks of plaque-purified progeny of the three different "phenotypes" was either labeled "in vivo" (left panel) or analyzed by Southern blot (17), using a labeled joint-spanning fragment (16) as probe (right panel). Three generations are shown in each panel (see the legend to Fig. 3 for explanation of labels). In the left panel, upper arrows point to the two forms of the joint fragment; lower arrows indicate S termini. In the right panel, only the S termini are highlighted by arrows in the *Bam*HI digests since the joint fragments did not resolve well. In the *Sac*I digest, arrows indicate the two joint fragments at the bottom of the "a" sequence stepladder as well as the S termini.

What is the mechanism by which a heterozygote segregates both classes of homozygous genome? We propose that this may occur by either or both of two pathways. The first possibility is that restoration of identity occurs by gene conversion during intramolecular recombination between the inverted "a" and "c" sequences (Fig. 7A). By this model, recombination events initiated at the "a" sequence resulted in regions of heteroduplexed DNA which can migrate into the flanking "c" sequences. Whenever there is sequence heterology, mismatch repair can occur (6, 10) either by looping out of unmatched sequences (Fig. 7A-i), resulting in a deletion, or by new synthesis across the unmatched sequence which is used as a template (Fig. 7A-ii). Repair may occur either on the spot or during the next round of viral DNA synthesis. This model proposes that although genome segment inversion is initiated at the "a" sequence (12), the actual site of recombination, or resolution of the Holliday structure, may occur anywhere within the potentially paired diploid sequences. Another mechanism by which homozygotes may be generated is intermolecular exchange between S segments that are inverted with respect to each other. There are several different combinations of crossover events that can result in exchange of the heterologous regions, some of which are depicted in Fig. 7B. This model requires that no constraints be placed on the initiation and resolution of recombination. If such constraints are placed on the model, elaborate repair mechanisms must be invoked to account for the results observed in this study. For example, if initiation and resolution of recombination do not involve exchange across the mismatched region, then crossovers alone can account for the data. Where exchange is not initiated between the heterologous regions within the inverted repeats, a second crossover must occur to result in homozygous progeny (bottom left panel). The panel on the right is a special case since the two genomes are misaligned. However, if exchange passes through the diploid region, then mismatch repair of the heteroduplex or subsequent viral DNA synthesis must be invoked. At present, it is impossible to distinguish between recombination and conversion, although it is worth stressing that both rely on similar mechanisms and they are not mutually exclusive. We suspect that

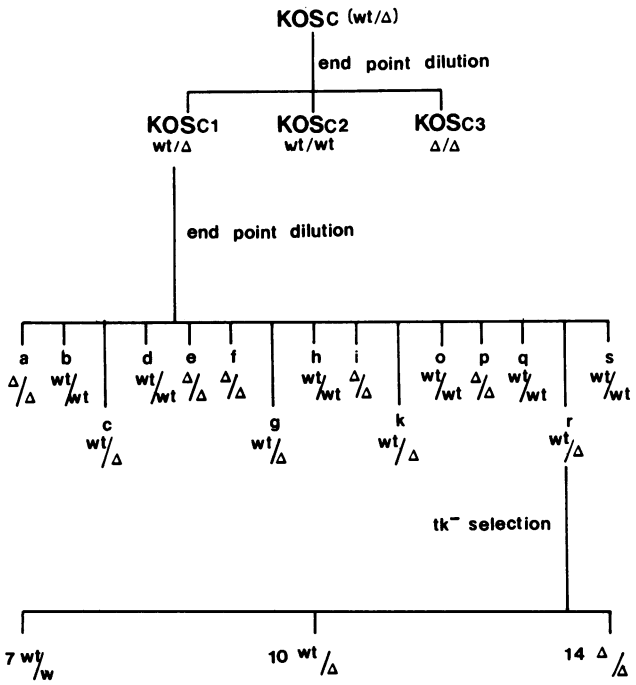

FIG. 5. Genealogy of KOSC progeny. The "phenotype" is indicated by wt/Δ (heterozygous), wt/wt (homozygous for the larger form of the "c" sequence), or Δ/Δ (homozygous for the smaller form of the "c" sequence). The letters in the KOSC1 progeny group are individual codes, and the numbers in the last group indicate the number of progeny with a particular "phenotype."

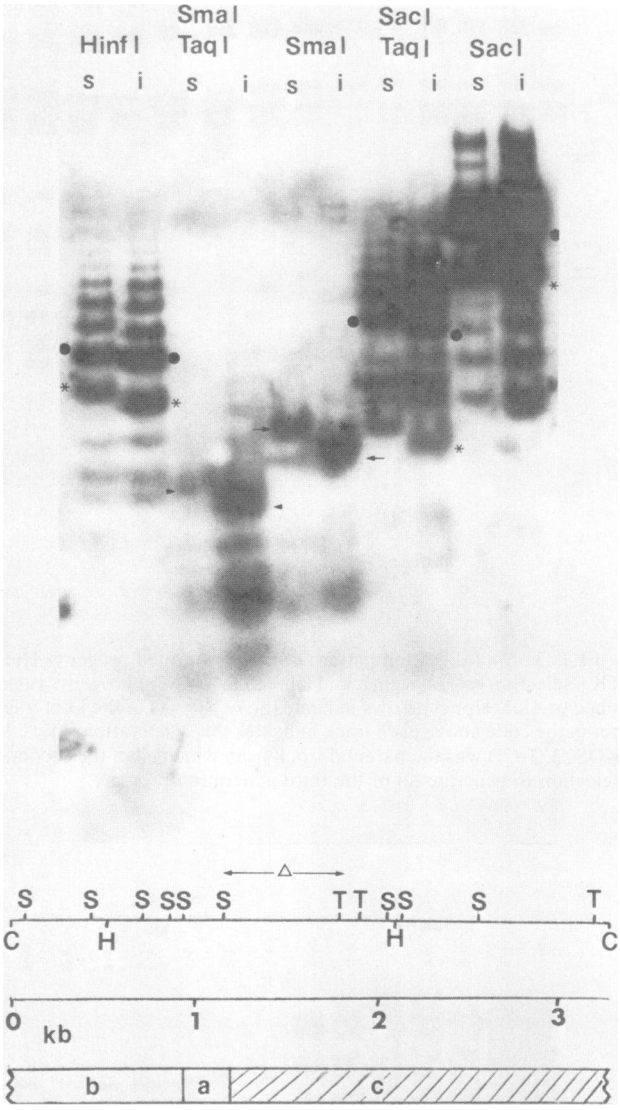

FIG. 6. Southern blot map of deletion. The DNAs from two homozygous progeny of KOSC1—KOSC1i (smaller form) and KOSC1s (larger form)—were digested with various restriction enzymes, electrophoresed on a 1.4% agarose gel, blotted to nitrocellulose as described by Southern (18), and hybridized with a cloned joint-spanning fragment (16). The difference between the two virus stocks was localized to the *Sma*I-*Taq*I fragment (indicated by arrowheads) which maps adjacent to the *Sma*I fragment containing most of the "a" sequences (see Fig. 1). Dots appear on the left of joint fragments in tracks marked "s" and on the right of joint fragments in tracks marked "i." Similarly, asterisks indicate S terminal fragments; arrows indicate the heterozygous *Sma*I fragment, and arrowheads indicate the heterozygous *Sma*I-*Taq*I fragments.

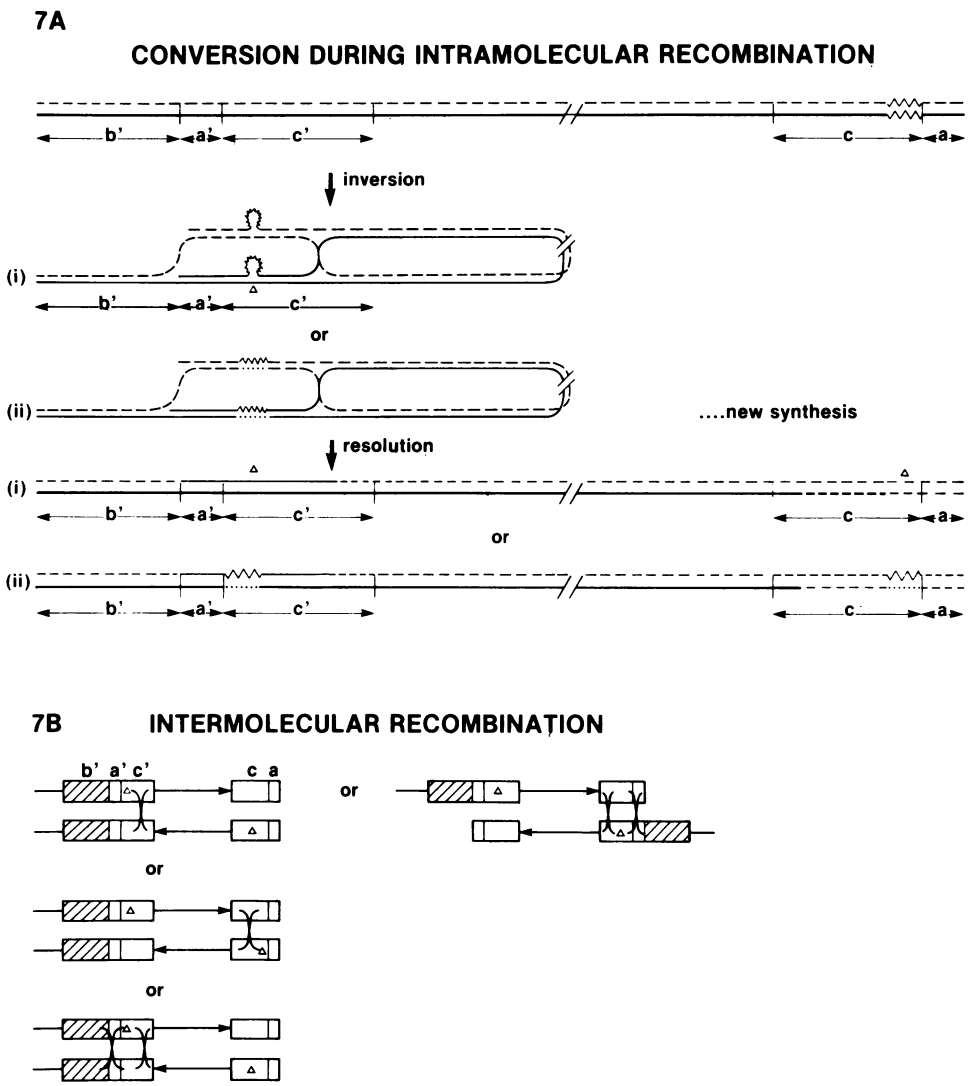

FIG. 7. (A) Conversion during intramolecular recombination. The two strands of viral DNA are distinguished by solid and dashed lines, and the zig-zag represents the extra 60 bp present in the larger form of the "c" sequence. The inverted repeats are indicated below and are not drawn to scale. (B) Intermolecular recombination. The inverted repeats flanking Us are depicted by open boxes. Only the S segment of the HSV genome is shown here. The symbol Δ represents the 60-bp deletion in the "c" sequence. χ indicates the location of possible crossover events.

in either case, the process of generating homozygotes depends ultimately on the high-frequency specialized recombination events occurring at the L-S junction, because both processes occur at comparably high frequencies. If this suggestion is correct, it would imply that recombination events initiated at the "a" sequence can, given a flanking region of homology, be propagated some distance before resolution. We are currently testing this hypothesis.

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute of Canada and the Medical Research Council of Canada. J. R. Smiley is a research scholar of the National Cancer Institute of Canada.

LITERATURE CITED

1. **Clements, J. B., J. McLauchlan, and D. McGeoch.** 1979. Orientation of herpes simplex virus type 1 immediate early mRNAs. Nucleic Acids Res. **7**:77–92.
2. **Davison, A., and N. Wilkie.** 1981. Nucleotide sequences of the joint between the L and S segments of herpes simplex virus types 1 and 2. J. Gen. Virol. **55**:315–331.
3. **Davison, A., and N. Wilkie.** 1983. Inversion of the two segments of the herpes simplex virus genome in intertypic recombinants. J. Gen. Virol. **64**:1–18.
4. **Hayward, G., N. Frenkel, and B. Roizman.** 1975. Anatomy of herpes simplex virus DNA: strain differences and heterogeneity in the locations of restriction endonuclease cleavage sites. Proc. Natl. Acad. Sci. U.S.A. **72**:1768–1772.
5. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
6. **Holliday, R.** 1964. A mechanism for gene conversion in fungi. Genet. Res. **5**:282–304.
7. **Knipe, D., W. Ruyechan, R. Honess, and B. Roizman.** 1979. Molecular genetics of herpes simplex virus. The terminal *A* sequences of the L and S components are obligatorily identical and constitute a part of a structural gene mapping predominantly in the S component. Proc. Natl. Acad. Sci. U.S.A. **76**:4534–4538.
8. **Knipe, D., W. Ruyechan, B. Roizman, and I. Halliburton.** 1978.

Molecular genetics of herpes simplex virus: demonstration of regions of obligatory and nonobligatory identity within diploid regions of the genome by sequence replacement and insertion. Proc. Natl. Acad. Sci. U.S.A. **75**:3896–3900.

9. **Machem, S., and B. Roizman.** 1980. Regulation of herpes virus macromolecular synthesis: transcription-initiation sites and domains of α genes. Proc. Natl. Acad. Sci. U.S.A. **77**:7122–7126.
10. **Messelson, M., and C. Radding.** 1975. A general model for genetic recombination. Proc. Natl. Acad. Sci. U.S.A. **72**:358–361.
11. **Mocarski, E., L. Post, and B. Roizman.** 1980. Molecular engineering of the herpes simplex virus genome: insertion of a second L-S junction into the genome causes additional genome inversions. Cell **23**:243–255.
12. **Mocarski, E., and B. Roizman.** 1982. Structure and role of the herpes simplex virus DNA termini in inversion circularization and generation of virion DNA. Cell **31**:89–97.
13. **Mocarski, E., and B. Roizman.** 1982. Herpes virus-dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral *A* sequences and linked to an origin of viral DNA replication. Proc. Natl. Acad. Sci. U.S.A. **79**:5626–5630.
14. **Schaffer, P. A., V. C. Carter, and M. C. Timbury.** Collaborative complementation study of temperature-sensitive mutants of herpes simplex virus types 1 and 2. J. Virol. **27**:490–504.
15. **Sheldrick, P., and N. Berthelot.** 1974. Inverted repetitions in the chromosomes of herpes simplex virus. Cold Spring Harbor Symp. Quant. Biol. **39**:667–678.
16. **Skare, J., and W. C. Summers.** 1977. Structure and function of herpes virus genomes II EcoRI, XbaI and HindIII endonuclease cleavage sites on herpes simplex virus type 1 DNA. Virology **76**:581–595.
17. **Smiley, J., B. Fong, and W.-C. Leung.** 1981. Construction of a double jointed herpes simplex virus DNA molecule: inverted repeats are required for segment inversion and direct repeats promote deletions. Virology **II**:345–362.
18. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
19. **Wagner, M., and W. C. Summers.** 1978. Structure of the joint region and the termini of the DNA of herpes simplex virus type 1. J. Virol. **27**:374–387.
20. **Watson, R., C. Preston, and J. B. Clements.** 1979. Separation and characterization of herpes simplex virus type 1 immediate-early mRNA's. J. Virol. **31**:42–52.
21. **Wilkie, N. M.** 1976. Physical maps for herpes simplex virus type 1 DNA for restriction endonucleases *Hind* III, *Hpa*-1, and *X. bad.* J. Virol. **20**:222–233.