JOURNAL OF VIROLOGY, June 1989, p. 2690–2698
0022-538X/89/062690-09$02.00/0

Vol. 63, No. 6

Nucleotide Sequences at Recombinational Junctions Present in Pseudorabies Virus Variants with an Invertible L Component

ZHIQIANG LU, JEAN M. DEMARCHI,† LINDA HARPER, GLENN F. RALL, AND TAMAR BEN-PORAT*

Department of Microbiology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received 3 January 1989/Accepted 8 March 1989

The genome of pseudorabies virus (PrV) consists of two components—a noninvertible long (L) and an invertible short (S) component. The S component is bracketed by inverted repeats. In some variant strains of PrV (which have a selective growth advantage in certain cell lines), a sequence normally present at the left end of the L component has been translocated to the right end of the L component next to the inverted repeat. Consequently, these strains have acquired a genome with an L component that is bracketed by inverted repeats and that also inverts. We have determined the restriction maps and have analyzed the nucleotide sequences of those parts of the genome of three strains with invertible L components that contain the translocated segment of DNA. The results were as follows. The translocated fragments were derived in all cases from the extreme left end of the L component only. The sizes of the translocated fragments were similar, ranging from 1.3 to 1.4 kilobase pairs. The junction between the L and S components in these strains was the same as that in standard viral concatemeric DNA. The translocation of sequences from the left end of the genome next to the inverted repeats was always accompanied by a deletion of sequences from the right end of the L component. The sizes of the deleted fragments varied considerably, ranging from 0.8 to 2.3 kilobase pairs. Sequence homology at the points of recombination, i.e., at the junction between the right end and the left end of the L component, existed sometimes but not always. A model depicting how a virus with a class 2 genome (such as PrV) may acquire a genome with characteristics of a class 3 genome (such as herpes simplex virus) is presented.

The herpesviruses constitute a large family of DNA viruses whose genomes consist of linear, double-stranded, noncircularly permuted DNA molecules. A particularly interesting feature of these molecules was first observed by Sheldrick and Berthelot (19), who reported that herpes simplex virus DNA contains inverted repeats (i.e., palindromic sequences bracketing sets of unique sequences). This genome organization was found thereafter to be common in many herpesviruses (1, 9, 18). On the basis of the presence of inverted or direct repeated sequences, the genomes of the herpesviruses have been divided into three (9) or five (18) classes.

The genome of pseudorabies virus (PrV) is a class 2 (9) or class D (18) DNA molecule and consists of two unique sequences, the unique long (U_L) and the unique short (U_S). The U_S is bracketed by inverted repeats and inverts itself relative to the U_L sequence. As a consequence, the genomes are found in two equimolar isomeric forms (4, 21). More than 40 field isolates of PrV have been examined, and even though the restriction patterns of the DNA of the various field strains vary somewhat (2, 6, 8, 14), all contain genomes in which the L component is fixed in only one orientation relative to the S component, whereas the S component inverts itself. However, several attenuated vaccine strains have genomes that behave like class 3 or class E DNA molecules (of which the genome of HSV is a prototype); i.e., both the U_S and the U_L components are bracketed by inverted repeats and both invert (11).

PrV variants that have acquired an invertible L component are, in general, obtained after extensive passage of the virus in cell culture, especially chicken embryo fibroblast (CEF) cell cultures; repeated passage of PrV in CEF, but not in rabbit kidney (RK) or pig kidney (PK) cells, results in the evolution of populations of virions with genomes that have an invertible L component (12). Thus, certain growth conditions appear to provide a selective pressure for the evolution of virions with these types of genomes. The invertibility of the L component in the genome of these PrV variants is associated with a translocation of sequences derived from the left end of the genome to a position next to the S component. The conversion of a standard PrV genome to a genome with an invertible L component is thus invariably linked to the emergence of an L component that has become bracketed by inverted repeats (12).

The functions of the inverted repeats in the genome of the herpesviruses, the functional significance of the inversions they promote, and the evolutionary pressures that have led to the emergence of herpesviruses with different class genomes are not clear. The evolution of PrV variants that have acquired an invertible L component as a result of passage in CEF allows one to identify the events leading to the conversion of a class 2 genome to one that has characteristics of a class 3 genome and may help in the understanding of the significance of these genome structures. To attempt to answer some aspects of these questions, we have analyzed the nucleotide sequences at the recombinational junctions bordering the newly formed inverted repeats in PrV variants that have acquired an invertible L component and have compared them with similar sequences present in the wild-type genome.

MATERIALS AND METHODS

Virus strains and cell culture. PrV(Ka) is a strain that has been carried in our laboratory for more than 25 years. The Norden strain is an attenuated vaccine strain; its origins have been described previously (13, 16). The origin and derivation of the other strains used are described in the text.

* Corresponding author.
† Present address: Department of Pathology, State University of New York Health Science Center, Syracuse, NY 13210.

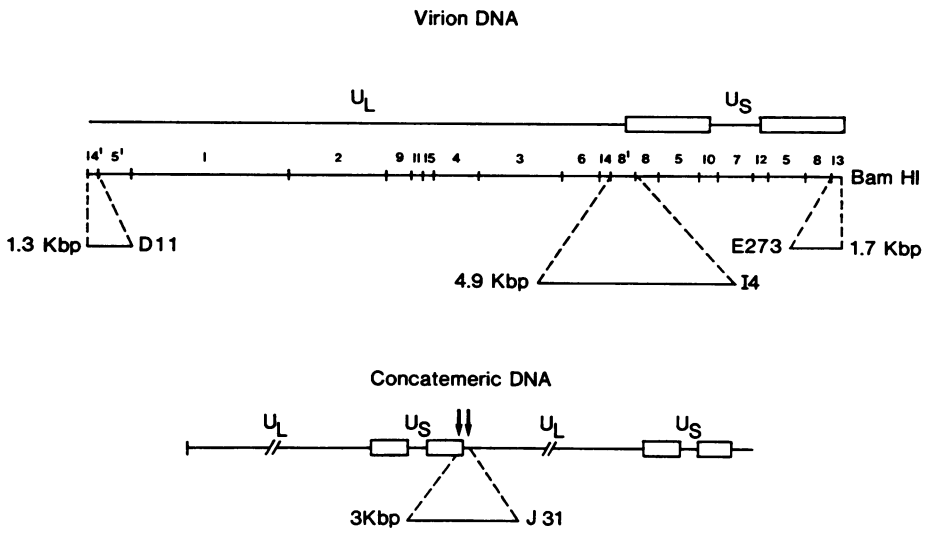

FIG. 1. Location of end fragments and junction fragments obtained from standard DNA. The junction fragment I4 contains the *Bam*HI junction between the L and S components of mature virion DNA. The junction fragment J31 contains the *Bam*HI junction fragments between the two ends of the genome which is found in intracellular concatemeric virus DNA; J31 contains the sequence of both end fragments of mature DNA, D11 and E273. ▭, Inverted repeats bracketing the U_S.

PK and CEF cells were cultivated in Eagle synthetic medium (5) containing 5% bovine serum.

Enzymes and chemicals. Restriction enzymes, DNA polymerase I, and T4 DNA ligase were purchased from Bethesda Research Laboratories, Inc. T4 polynucleotide kinase was purchased from United States Biochemical Corp. [γ-^{32}P]ATP and [α-^{32}P]dCTP were purchased from Dupont, NEN Research Products.

Purification of virions and of viral DNA. A continuous line of PK cells was infected (multiplicity of infection, 0.5 PFU per cell) and incubated for 48 h in Eagle medium at 37°C. Virions were purified as described previously (3) and were lysed by the addition of sarcosyl (final concentration, 2%). The samples were incubated for 15 min at 60°C and digested with nuclease-free pronase (1 mg/ml) for 2 h. The DNA was extracted with phenol-chloroform-isoamyl alcohol (50:48;2, vol/vol/vol) and dialyzed against buffer (0.01 M Tris, 0.001 M EDTA [pH 7.6]).

Cloning of restriction fragments. The *Bam*HI restriction fragments of interest were cloned into the *Bam*HI site of pBR325 as described previously (7, 10). Appropriate subfragments were isolated, blunt ended with Klenow, and cloned into the *Sma*I site of pUC13.

Southern analysis. Digested DNA fragments were separated on agarose gels and transferred to nitrocellulose filter paper (20). The immobilized DNA bands were hybridized to [α-^{32}P]dCTP nick-translated DNA probes (17).

DNA sequencing. DNA was sequenced by the method of Maxam and Gilbert (15).

RESULTS

Origin of the variants of PrV that have acquired an invertible L component and cloning of the junction fragments between the L and S components. The junction fragments between the L and S components of four different PrV variants were cloned and mapped.

(i) PrV(Ka) strain. The PrV(Ka) strain is a standard laboratory strain with a noninvertible L component. The ends of the mature DNA, as well as the junction fragments between the L and S components in the mature genome and in concatemeric DNA, were cloned into the *Bam*HI site of pBR325 (7). Clone D11 contains the left end of the genome (*Bam*HI fragment 14'); clone E273 contains the right end of the genome (*Bam*HI fragment 13); clone I4 contains the normal junction fragment between the S and L components present in the mature genome (*Bam*HI fragment 8'); clone J31 contains the *Bam*HI junction fragment obtained from concatemeric DNA (*Bam*HI fragment 14' + *Bam*HI fragment 13). The locations of these *Bam*HI fragments on the genome of PrV(Ka) are illustrated in Fig. 1.

(ii) PrV(Ka)B8 variant. The PrV(Ka)B8 variant is a plaque isolate derived from a population of standard PrV(Ka) virions that had been passaged 30 times in CEF. In this variant, sequences of DNA normally only present at the left end of the U_L in the genome of standard virus have been translocated next to the inverted repeat (12). As a consequence, the U_L component of the genome inverts itself relative to the U_S component. Clone B8/130 contains the junction fragment between the L and S components in the standard orientation of the L component. It is equivalent to *Bam*HI fragment 8' of the standard PrV genome but contains a segment of DNA derived from the left end of the genome which has been translocated next to the internal inverted repeat of the S component (Fig. 2).

(iii) Norden strain. The Norden strain is a vaccine strain which also has experienced a translocation of sequences from the left end of the genome next to the inverted repeat and which has an invertible L component. *Bam*HI restriction fragments containing the junction fragments between the L and S components in the standard and the inverted orientation of the L component were cloned. Clone Norden 78 contains the L-S junction in the standard orientation of the L component; i.e., it is equivalent to *Bam*HI fragment 8' of the standard PrV genome. However, it includes a segment of DNA derived from the left end of the genome that had been translocated next to the internal inverted repeat. Clone Norden 75 contains the junction fragment derived from a genome in which the L component is inverted. This fragment

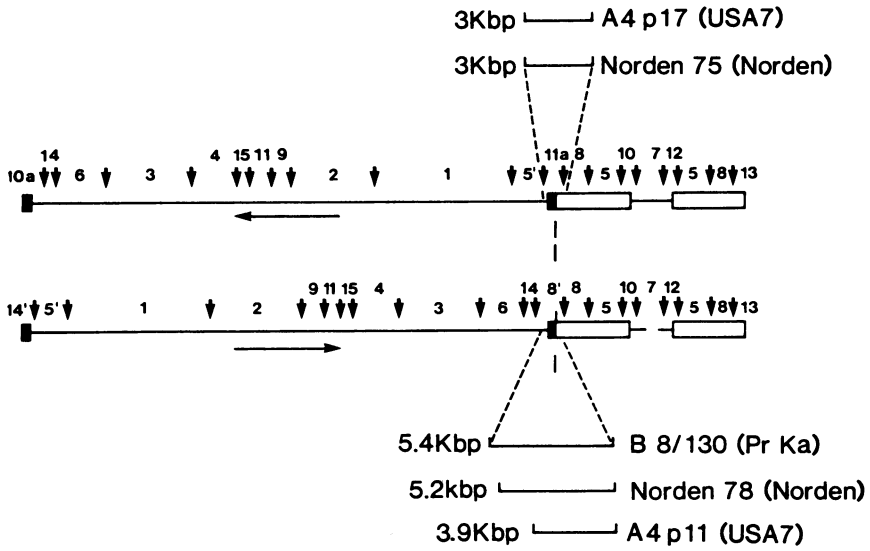

FIG. 2. Location of junction fragments obtained from strains of PrV with an invertible L component. The *Bam*HI junction fragments betwen the L and S components were cloned from the appropriate virion DNA. Since the L component inverts itself relative to the S component, two junction fragments from each variant can be obtained. B8/130 [PrV(Ka)] is the cloned *Bam*HI fragment derived from the prototype orientation of the L component; the fragment derived from the inverted orientation of the L component of B8/130 is equivalent to J31 of standard PrV(Ka) and has not been analyzed in detail. Norden 78 is the cloned *Bam*HI fragment derived from the prototype genome orientation of the L component of the Norden strain; Norden 75 is derived from the inverted orientation of the L component of this strain. A4(p11) is the cloned *Bam*HI fragment obtained from the prototype genome orientation of USA7(A4); A4p17 is the cloned *Bam*HI fragment obtained from the inverted orientation of the L component of this strain. ▼, *Bam*HI cleavage sites; →, orientations of the L component; ▭, inverted repeats bracketing the U_S; ▮, inverted repeats bracketing the U_L.

is equivalent to the junction fragment found in concatemeric DNA; i.e., it consists of the joined *Bam*HI 14'/*Bam*HI 13 end fragments (Fig. 1).

(iv) USA7A4. USA7A4 is a variant of a virus isolated in the United States, USA7. The USA7 strain was passaged 30 times in CEF and has, as a consequence, acquired an invertible L component (12). USA7A4 is a plaque isolate from this population of virions. Two *Bam*HI restriction fragments containing the junctions between the L and S components in the standard and inverted orientation of the L component were cloned. Clone A4P11 contains the L-S junction in the standard orientation of the L component; i.e., it is equivalent to *Bam*HI fragment 8' but includes a segment of DNA derived from the left end of the genome that had been translocated next to the internal inverted repeat; clone A4p17 contains the junction derived from the inverted form of the L component and is equivalent to the junction found in concatemeric DNA; it consists of the joined *Bam*HI end fragments.

Restriction mapping of the junction fragments. The cloned junction and end fragments of wild-type PrV(Ka) and of the variants were mapped with various restriction enzymes (Fig. 3). Lines A to D show the restriction maps of the clones obtained from standard PrV(Ka). As expected, the right side (derived from the S component) of the concatemeric junction (clone J31) and of the normal junction (I4) between the L and S components, as well as the right end of the genome (clone E273), are identical to each other. Also, as expected, the left end of the genome (clone D11) is identical to the left side of the junction fragment derived from concatemeric DNA (clone J31). The two ends of the L component differ, however.

Restriction mapping of the modified prototype junction of PrV(Ka)B8, the variant of PrV(Ka) that had acquired an invertible L component (Fig. 3, line E), indicates that a segment of DNA derived from the left end of the genome, which includes the *Nco*I site present approximately 200 nucleotides, but not the Sma I site present 20 nucleotides, from the left end of *Bam*HI fragment 14' (clone D11, line B), had been translocated to the right side of the L component. A deletion of the sequences from the right side of the L component up to approximately 700 nucleotides to the right of the *Kpn*I site was also observed. Thus, a recombinational event between sequences in *Bam*HI fragment 14' (somewhere within the 180 nucleotides between the *Nco*I site and the *Sma*I site) and a sequence approximately 0.7 kilobase pairs (kbp) to the right of the *Kpn*I site in *Bam*HI fragment 8' had occurred. This recombinational event resulted in approximately 1.3 kbp of the left end of the genome being translocated to the right side of the L component and approximately 0.8 kbp of DNA from the right side of the L component being deleted.

Figure 3, lines F and G, illustrates the restriction maps of the two junction fragments (in both orientations of the L component) obtained from Norden DNA. Again, as expected, the restriction patterns of the segments of DNA derived from the S segment, i.e., from the inverted repeats, are similar in both fragments. Comparison of the prototype junction of the Norden genome (clone Norden 78, line G) with that of PrV(Ka) (clone I4, line D) shows that the left sides of the maps are the same, starting at the unique *Kpn*I site. On the other hand, the right side of the L component of Norden 78 is similar to that found in the concatemeric junction of PrV(Ka) (clone J31, line C); i.e., it appears to be derived from the left end of the L component. Indeed, the restriction patterns of the 1.3 kbp of the L components adjacent to the S component are similar in Norden 75 and Norden 78, with the exception that the *Sma*I site present within the leftmost 20 nucleotides of Norden 75 is not present in Norden 78. These findings indicate that in the

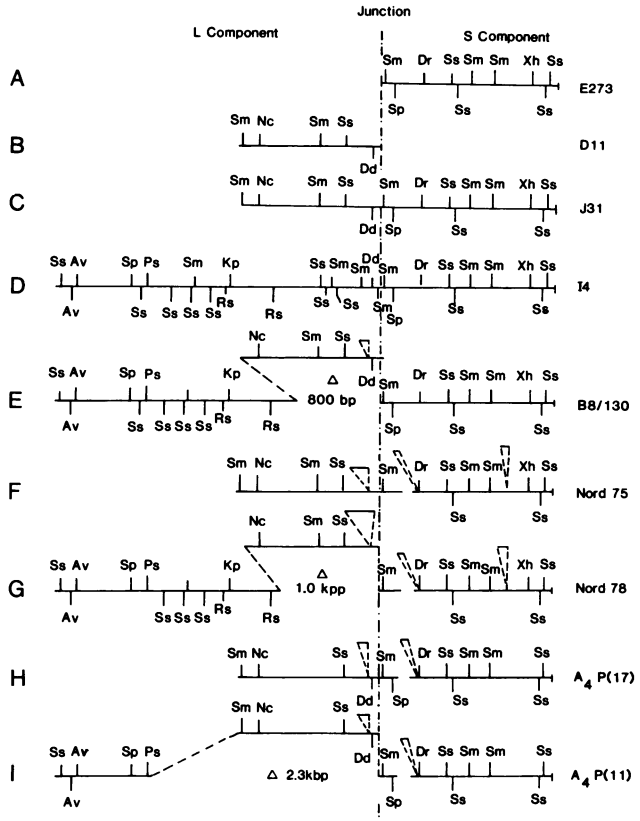

FIG. 3. Restriction maps of the end and junction fragments obtained from different PrV variants. The cloned *Bam*HI restriction fragments, the locations of which are shown in Fig. 1 and 2, were mapped with restriction enzymes, and the restriction maps were aligned. The segments of DNA that were inserted into the prototype junction of the variants with an invertible L component are raised. Deletions or insertions in the DNA fragments relative to those of PrV(Ka) are indicated by discontinuities or insertions in the horizontal lines. Ss, *Sst*II; Av, *Ava*II; Sp, *Sph*I; Ps, *Pst*I; Sm, *Sma*I; Nc, *Nco*I; Dr, *Dra*I; Dd, *Dde*I; Xh, *Xho*I; Kp, *Kpn*I.

Norden strain most of the sequences of *Bam*HI fragment 14′ (approximately 1.3 kbp of the leftmost sequences of the PrV genome) have been translocated next to the inverted repeat. The rightmost 1.0 kbp of the L component, however, was deleted. Thus, a recombinational event occurred between sequences in *Bam*HI fragment 14′ (somewhere between the *Nco*I site and the *Sma*I site) and a sequence within *Bam*HI fragment 8′ approximately 500 bp to the right of the *Kpn*I site.

Figure 3 (lines H and I) shows the restriction maps of the junction fragments between the L and S components obtained from USA7A4 DNA. In this case, most of *Bam*HI fragment 14′, including the leftmost *Sma*I fragment (and possibly more), was translocated next to the inverted repeat. On the other hand, a much larger part of the right side of the L component was deleted. The deleted sequence spans the right side of the L component up to somewhere between the *Pst*I site and the first *Sst*II site to the right of the *Pst*I site. This deletion is approximately 2.3 kbp in size.

Differences of up to approximately 300 nucleotides were also found in the different clones between the *Dde* site and the first *Sst*II cleavage site in the sequences normally found at the left end of the L component (i.e., near the end of the genome). This is due to the presence of a sequence of 34 nucleotides which is repeated 3 times in J31, 5 times in B8/130, 12 times in Norden 78, and 9 times in Norden 75. Other differences in the numbers of some direct repeats present within the S component were also observed in the different clones; these are indicated in Fig. 3 either as small deletions or as insertions.

Southern analysis of the junction fragments. The results of the restriction enzyme mapping of the junction fragments obtained from virus variants with genomes with invertible L components indicate that in all cases, between 1.3 and 1.4 kbp derived from the left end of the genome was translocated next to the inverted repeat. Furthermore, between approximately 0.8 and 2.3 kbp usually present at the end of the right side of the L component was deleted. Because the sequences derived from the left or right end of the L component of standard PrV(Ka) DNA do not cross-hybridize, the crossover points at which recombination between the two ends of the L component occurred can be ascertained by Southern hybridization. Using sequences derived from the two ends of the L component, as well as sequences derived from the end of the S component as probes, we localized the junction points at which recombination had occurred (data not shown). The results showed that the conclusions that we had drawn from the restriction mapping data (Fig. 3) regarding the location of these recombination points were correct.

Sequence analysis of the junction fragment. The translocation of the leftmost sequences of the L component next to the inverted repeat in the genome of variants which had acquired an invertible L component gives rise to two new junctions: (i) the junction between the internal inverted repeat and the translocated fragment—the R junction, and (ii) the junction between sequences near the right side of the L component and those of the translocated fragment—the L junction (Fig. 4). We have analyzed the nucleotide sequence at both junctions. The aim of these studies was to ascertain whether the R junction between the S and L component was the same as the junction found in concatemeric DNA. Also, we wanted to determine whether any homologies existed at the point of recombination between the right and left sides of the L component which might have facilitated the translocation event. The information concerning the regions in which the recombinational events had probably occurred and which should be sequenced was derived from the mapping data (Fig. 3).

The segments of DNA to be sequenced either were subcloned (as described in Materials and Methods) or were obtained by elution from agarose gels. They were mapped with restriction enzymes and sequenced. Several hundred nucleotides on each side of the junction fragments were sequenced. The sequence analysis showed that the fragment of DNA derived from the left end of the L component that was translocated to the right side of the L component (next to the inverted repeat) had not undergone any significant modifications; only an occasional nucleotide insertion or substitution was observed. Similarly, the nucleotide sequence at the right side of the L component was the same as that normally present at that position in standard virus, but, as mentioned above, a segment of DNA (which varied in size) normally present at the end of the right side of the L component was invariably deleted from the genomes of the variants that had acquired an invertible L component.

(i) Nucleotide sequence at the junction between the L and S components (R junction). Figure 5 (lines 1 and 2) shows the nucleotide sequence at the concatemeric junction of PrV(Ka) DNA, i.e., the junction between the two ends of the mature genome (clone J31), as well as the R junction in the

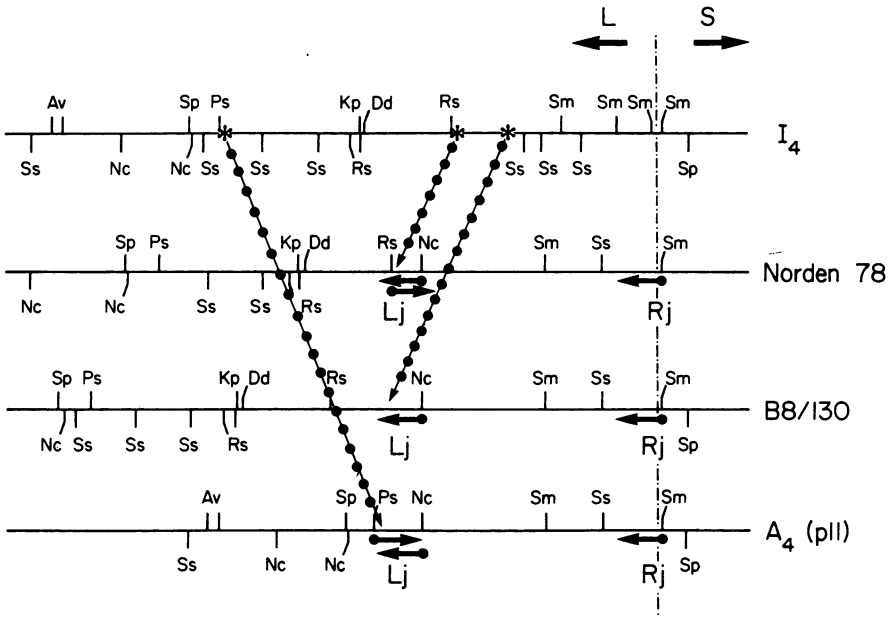

FIG. 4. Sequencing strategy used to analyze the recombinational junctions between the sequences from the left side of the L component that have been translocated and those at the right side of the L component. The horizontal arrows under the restriction maps indicate those segments of DNA in which the recombinational junctions have been located. Rj, Right junction (adjacent to the S component); Lj, left junction (the recombinational junction between the translocated fragment derived from the left end of the L component and the right side of the L component). The dotted lines indicate the points at which recombination between the two sides of the L component have taken place.

prototype arrangement of the genome of the invertible PrV(Ka)B8 variant. The nucleotide sequence at the two junctions is similar. The nucleotide sequences at the R junction in the prototypic or inverted orientations of the Norden and USA7A4 genomes were also almost identical to that of the concatemeric junction of PrV(Ka); only an occasional substitution or insertion of a nucleotide was observed. Furthermore, the nucleotide sequences at the R junctions of the prototype and the inverted orientation of the L components in the genomes of either the Norden or the USA7A4 strains were identical to each other.

The sequence at the junctions between the S and L components in the prototype orientation of the L components of genomes (with the translocation of the sequence derived from the left side of the L component to the right side of the L component, next to the inverted repeat) is in each case similar to that found in the concatemeric junctions. Therefore, it appears that the events that caused the translocation must have occurred between molecules in which the junction between the two ends of the genome had already formed, i.e., after the mature genome had either circularized or had become concatemeric.

(ii) Nucleotide sequence at the recombination point between the right and left side of the L component (L junction). (a) The Norden strain. The nucleotide sequence at the L junction, i.e., the recombination points between the right and left sides of the L component in the Norden strain, is illustrated in Fig. 6. The junction was found in the position predicted from the restriction mapping data. Almost the entire *Bam*HI 14′ fragment, with the exception of the last 70 nucleotides of this fragment (1,340 bp), had been translocated next to the S component. This sequence abuts a sequence 994 bp inward

	R-JUNCTION				
	L COMPONENT			S COMPONENT	
J31	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC	CCCACCCCCC	GCTCCCCGGG GGCCGCGAA.
B8/130	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC (A/T)	CCCACCCCCC	GCTCCCCGGG GGCCGCGAAA
NORDEN 75	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC	TCGGGCCCCC	GCTCCCCGGG GGCCGCGAAA
NORDEN 78	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC	TCGGGCCCCC	GCTCCCCGGG GGCCGCGAAA
A4(P17)	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC	CCCACCCCCC	GCTCCCCGGG GGCCGCGAAA
A4 (P11)	AGAGATCCGC	GCTCGGGGGA	GAGCTGGGCC	CCCACCCCCC	GCTCCCCGGG GGCCGCGAAA

FIG. 5. Nucleotide sequence at the R junction of different variants. The nucleotide sequences were obtained by sequencing from the *Sma*I site present in the S component, 16 nucleotides upstream from the L-S junction (Fig. 4). Boldface letters indicate differences between the sequence in the variants and the corresponding sequence found in concatemeric DNA of PrV(Ka) (J31).

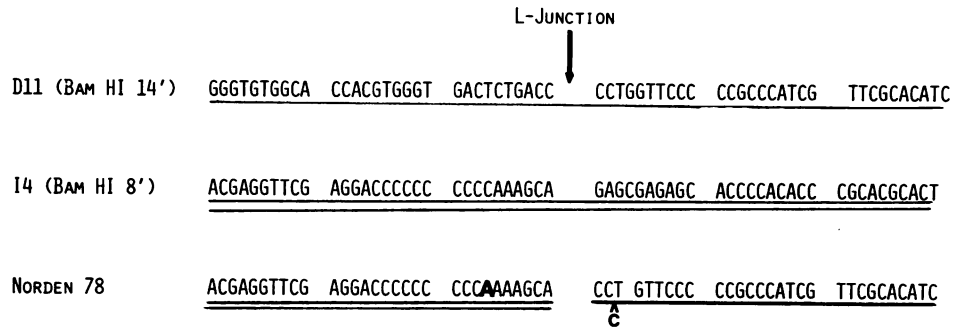

FIG. 6. Nucleotide sequence at the L junction of Norden 78. The nucleotide sequences of *Bam*HI fragment 14' (clone D11) and of *Bam*HI fragment 8' (clone I4) are available (G. Rall, J. DeMarchi, Z. Lu, and T. Ben-Porat, manuscript in preparation). The sequence of the L junction of Norden 78 was obtained by sequencing rightward from the *Rsa*I site and leftward from the *Nco*I site (Fig. 4). Sequences in Norden 78 derived from *Bam*HI fragment 14' are underlined once; those derived from *Bam*HI fragment 8' are underlined twice. Boldface letters indicate differences in the sequences in the variants from the corresponding sequences present in *Bam*HI fragment 14' or 8' of PrV(Ka).

of the right end of the L component (47 nucleotides to the right of the *Rsa*I site [Fig. 4]). Thus, as predicted from the mapping data, the rightmost approximately 1.0 kbp of the L component was deleted. No significant sequence homology near the recombination points that might have facilitated the recombinational event leading to the translocation was detected (Fig. 6).

(b) Clone B8/130 of PrV(Ka)B8. In this case also (Fig. 7), almost the entire *Bam*HI 14' fragment, with the exception of the last 50 nucleotides (1,360 bp), was translocated next to the S component. This sequence abuts a sequence 113 nucleotides to the left of the *Sst*II site (Fig. 4); i.e., the rightmost 798 bp of the L component was deleted. A sequence homology of 7 nucleotides was found around the point at which recombination had taken place (GGCGGGTG in D11 and GGCGGTG in I4).

(c) Clone USA7B4 (P11). The nucleotide analysis of the L junction in USA7B4 (Fig. 8) showed that the entire *Bam*HI 14' fragment (including part of the *Bam*HI site) had been translocated next to the S component (1,409 nucleotides). The translocated sequence abuts a sequence 2.3 kbp inward from the right end of the L component (23 nucleotides from the *Pst*I site [Fig. 4]); i.e., the rightmost 2.3 kbp of the L component has been deleted. A sequence homology of 6 nucleotides was found at the point at which recombination had taken place (GAGGATC in D11 and GAGATC in I4).

DISCUSSION

We have analyzed the nucleotide sequences spanning the junctions of the S and L components in the genomes of PrV variants that had acquired an invertible L component. Similarly, we have mapped and sequenced approximately 1,500 nucleotides at each of the ends of the genome of standard PrV, as well as at the junctions between the L and the S components present in mature and in concatemeric viral DNA. Because the presentation of these data is cumbersome and is not particularly informative in the context of the subject discussed in this article, we have confined ourselves to presenting only the nucleotide sequences near the junctions between the L and S components. We have also presented the junctions at the recombination points between the segment of DNA from the right and left parts of the L component resulting from the translocation of sequences from the left end to the right end of the L component in variants containing genomes with an invertible L component. More extensive sequencing data will be published elsewhere (J. M. DeMarchi, G. Rall, Z. Lu, and T. Ben-Porat, manuscript in preparation). Although we did not furnish in this paper all the available nucleotide sequences, it should be mentioned that the data showed that the segments of DNA that had been translocated to the right of the L component in the variants with an invertible L component had not undergone extensive modification and were in all cases derived solely from sequences normally present at the left end of the L component (the left end of the genome).

In the three independently isolated variants that we have analyzed that had acquired an invertible L component, the sizes of the translocated fragment were very similar—1.340, 1.360, and 1.409 kbp; i.e., they varied by only 69 nucleotides. Also, in all three variants the translocation of the

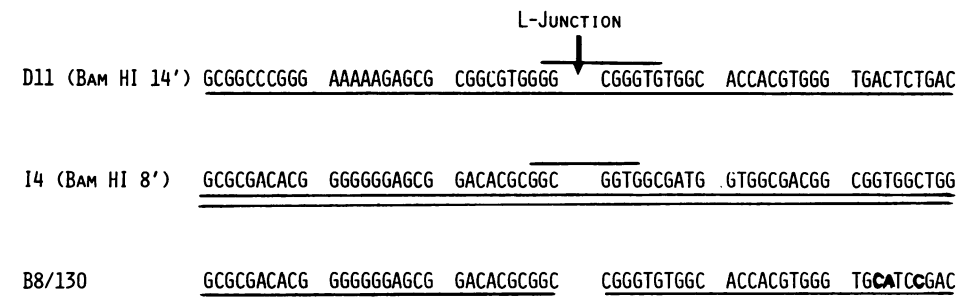

FIG. 7. Nucleotide sequence at the L junction of B8/130. The sequence of the L junction of B8/130 was obtained by sequencing leftward from the *Nco*I site (Fig. 4). Sequence homologies are indicated by a line over the nucleotides. Sequences in B8/130 derived from *Bam*HI fragment 14' are underlined once; those derived from *Bam*HI fragment 8' are underlined twice. Boldface letters indicate differences in the sequences in the variants from the corresponding sequences present in *Bam*HI fragment 14' or 8' of PrV(Ka).

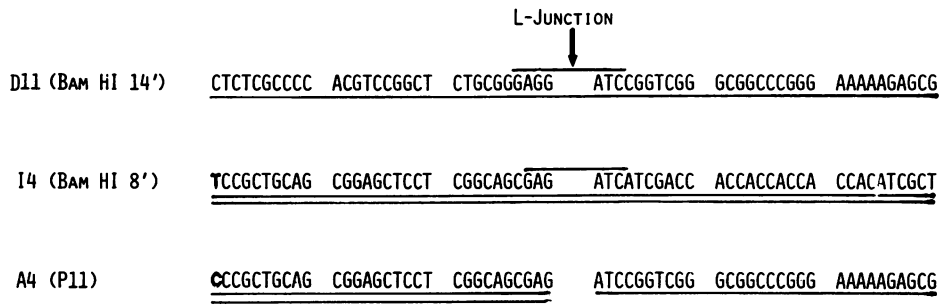

FIG. 8. Nucleotide sequence at the L junction of A4(p11). The sequence of the L junction of A4(p11) was obtained by sequencing leftward from the *Nco*I site and rightward from the *Pst*I site (Fig. 4). Sequence homologies are indicated by a line over the nucleotides. Sequences in A4(p11) derived from *Bam*HI fragment 14′ are underlined once; those derived from *Bam*HI fragment 8′ are underlined twice. Boldface letters indicate differences in the sequences in the variants from the corresponding sequences present in *Bam*HI fragment 14′ or 8′ of PrV(Ka).

sequences from the left end of the genome to the right end of the L component was accompanied by a deletion of sequences normally present in wild-type virus at the right end of the L component. However, in contrast to the sizes of the translocated segments of DNA, which were similar, the sizes of the deleted segments of DNA differed significantly in the different variants (between 0.8 and 2.3 kbp). The nucleotide sequences of the segments of DNA that remained at the right side of the L component and that were now adjacent to the translocated sequences remained unchanged in all cases.

The sequence analysis of the junction between the right and the left ends of the L component that had been translocated (the L junction) revealed a limited sequence homology near the points of recombination in some but not all cases. Sequence homology at the recombination point is thus not necessary for the recombination leading to the translocation of sequences from the left end of the L component to the right end of the L component.

The sequence analysis of the junction between the internal inverted repeat and the right end of the L component (the R junction) in the invertible strains revealed that these junctions are identical in the prototypic and inverted orientation of the L components of the Norden and USA7A4 strains. Furthermore, in PrV(Ka)B8 the nucleotide sequence at the junction of the prototypic arrangement of the L component is similar to that found at the junction of concatemeric DNA of standard PrV(Ka). These findings suggest that the translocation of the sequences derived from the left end of the L component to the right end of the L component was due to recombinational events involving concatemeric or circular DNA.

Figure 9 is a model showing how a genome could be generated in which the L component has become bracketed by inverted repeats derived from sequences normally only present at the left end of the L component. Standard PrV has a class 2 genome in which only the S component is bracketed by inverted repeats. Recombination between genomes in inverse orientation can therefore occur via the inverted repeats bracketing the S component (Fig. 9). This first recombinational event (R1) would give rise to DNA in which the cleavage-encapsidation signals would be separated by approximately 150×10^6 daltons of DNA rather than by 90×10^6 daltons, the size of standard DNA. While the capsid of PrV can accommodate a variable amount of DNA, molecules almost twice the size of the standard genome would probably be too large to be encapsidated, and this genome would therefore not be viable. To be viable, a second recombinational event (R2) would be necessary. In principle, this second recombinational event could occur in the inverted repeat, thereby giving rise to a normal genome; i.e., neither recombinational event would be detectable. Alternatively, the second recombinational event could occur in any part of the genome, for instance, in any part of the L component. However, the section of the L component between the inverted repeat and the recombination site would be deleted. A viable genome would be created only if a nonessential part of the L component was deleted, i.e., if the second recombinational event occurred at any position near the right end of the L component that is nonessential for growth. The rightmost 2.3 kbp of the L component is not essential for growth in vitro; it is deleted from USA7A4.

The model in Fig. 9 shows how a double recombinational event (or any even number of recombinational events) between concatemeric (or circular) DNA molecules in opposite orientation would generate genomes in which a segment of DNA derived from the left end of the L component would be translocated to the right end of the L component and sequences normally present at the right end of the L component would concomitantly be deleted. One of the recombinational events between molecules in opposite orientation would take place within the inverted repeats, and the second such event would take place within the right side of the L component, thereby eliminating a region of DNA in the L component that is nonessential for growth in vitro. This model is consistent with all our findings. In particular, it explains why the nucleotide sequences at the R junction in the genomes with the translocations and in concatemeric DNA obtained from standard genomes are identical. Furthermore, it explains why concomitantly with the translocation of sequences from the left side of the L component, deletions of varying sizes have occurred at the right side of the L component.

It is interesting that in the variants of PrV with an invertible L component, the inverted repeats that bracket that L component are derived exclusively from sequences normally present at the left end of the L component; no sequences derived from the S component are present in the repeats bracketing the L component. Since in standard PrV only the S component is invertible and no significant sequence homology exists between the termini of the L and S components (1,500 bp have been analyzed at each of the ends [unpublished results]), it appears that the inversion of the L component in the variants of PrV in which the L component has become bracketed by inverted repeats is not mediated by sequence-specific proteins. It has been shown that sequences derived from both ends of the mature viral genomes are required for efficient cleavage and encapsidation of concatemeric DNA (22). In the invertible variants,

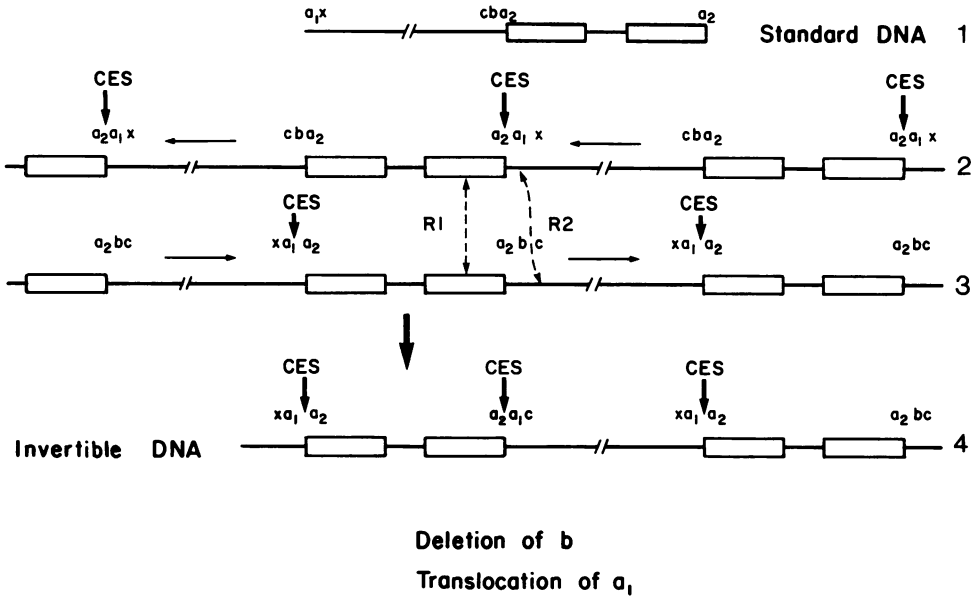

FIG. 9. Model showing how genomes with invertible L components can arise. Line 1 illustrates the structure of standard viral DNA. In lines 2 and 3, recombination between molecules of concatemeric standard DNA in opposite orientations is depicted. Two recombinational events, one in the inverted repeat (R1) and one in the L component (R2) (or any even number of recombinational events, half in the inverted repeat and half in the L component), will give rise to a genome in which sequences normally present at the left end of the genome (sequence a_1) are translocated to the right end of the L component. A concomitant deletion of a sequence normally present at the right end of the L component (sequence b) will also occur. Viable genomes will be formed only if the sequence that is deleted is not essential under certain growth conditions. CES, Cleavage-encapsidation sites.

the translocation of sequences derived from the left end of the genome next to the internal inverted repeat creates a cleavage-encapsidation site at the modified junction of the L and S components. It is noteworthy that cleavage at this newly formed junction would result in the formation of genomes in which the L component is inverted relative to the S component. Thus, the invertibility of the L component in these variants could be due either to the presence of two alternative sites for cleavage and encapsidation or to general recombination between the inverted repeats bracketing the L component.

PrV strains with genomes that have acquired an invertible U_L, because of the translocation of sequences derived from the left end of the L component to the right end of the L component, have a selective growth advantage over standard PrV under certain growth conditions (12). Indeed, the USA7A4 and PrV(Ka)B8 variants that we have analyzed were selected by passage in CEF, in which they have a growth advantage. The Norden vaccine strain has also been passaged extensively in embryonated eggs and CEF. The fact that the sizes of the segments of DNA that have been translocated are very similar in the three cases that we have analyzed suggests that the juxtaposition of these sequences next to the S component in the linear mature virus genome may be necessary to provide the observed growth advantage in CEF. Indeed, as mentioned above, sequences derived from both ends of the genome are required for efficient cleavage and encapsidation of concatemeric DNA (22). Concatemeric DNA of PrV with the translocation therefore has two cleavage-encapsidation sites per mature-genome-size DNA instead of the single cleavage-encapsidation site present in concatemeric DNA of standard PrV. It is possible that the extra cleavage encapsidation site provides a growth advantage to the virus under certain conditions. On the other hand, since a deletion of the rightmost sequences of the L component always accompanies the translocation of sequences from the left to the right side of the L component, the deletion could also possibly be responsible for the growth advantage of the virus in CEF. Our preliminary evidence suggests, however, that it is the juxtaposition of the sequences derived from the left end of the L component next to the inverted repeat in the mature genome that is responsible, at least in part, for the growth advantage the variant genomes have in CEF (G. Rall and T. Ben-Porat, unpublished results). We are currently investigating this problem in greater detail.

The evolutionary pressures that have led to the emergence of herpesviruses with different class genomes are obscure. The observations that PrV, which has a genome that is a class 2 molecule, may acquire a genome with characteristics of a class 3 molecule and that virions with these variant genomes have a growth advantage under certain conditions provide an initial insight into the possible evolutionary significance of the invertible elements present in the genomes of the herpesviruses.

ACKNOWLEDGMENT

This investigation was supported by Public Health Service grant AI-10947 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Ben-Porat, T.** 1982. Organization and replication of herpesvirus DNA, p. 147–172. *In* A. S. Kaplan (ed.), Replication of viral DNA. CRC Press, Inc., Boca Raton, Fla.
2. **Ben-Porat, T., A. M. Deatly, B. C. Easterday, D. Galloway, A. S. Kaplan, and S. McGregor.** 1983. Latency of pseudorabies virus, p. 365–383. *In* G. Wittman, R. M. Gaskell, H.-J. Rziha (ed.), Latent herpes virus infections in veterinary medicine. Martin Nijhoff, Boston.

3. **Ben-Porat, T., J. M. DeMarchi, and A. S. Kaplan.** 1974. Characterization of defective interfering viral particles present in a population of pseudorabies virions. Virology **61**:29–37.
4. **Ben-Porat, T., F. J. Rixon, and M. Blankenship.** 1979. Analysis of the structure of the genome of pseudorabies virus. Virology **95**:285–294.
5. **Eagle, H.** 1959. Amino acid metabolism in mammalian cell cultures. Science **130**:432–437.
6. **Gielkens, A. L. J., and A. J. M. Berns.** 1982. Differentiation of Aujeszky's disease virus strains by restriction endonuclease analysis of the viral DNAs. Curr. Top. Vet. Med. Anim. Sci. **17**:3–13.
7. **Harper, L., J. DeMarchi, and T. Ben-Porat.** 1986. Sequence of the ends of the genome and of the junction between the ends in concatemeric DNA of pseudorabies virus. J. Virol. **60**:1183–1185.
8. **Heppner, B., G. Darai, B. Podesta, G. Pauli, and H. Ludwig.** 1981. Strain differences in pseudorabies viruses. Med. Microbiol. Immunol. **169**:128.
9. **Honess, R. W., and D. H. Watson.** 1977. Unity and diversity in the herpesviruses. J. Gen. Virol. **37**:15–37.
10. **Ladin, B. F., S. Ihara, H. Hampl, and T. Ben-Porat.** 1982. Pathway of assembly of herpesvirus capsids. An analysis using DNA^+ temperature sensitive mutants of pseudorabies virus. Virology **116**:544–561.
11. **Lomniczi, B., M. L. Blankenship, and T. Ben-Porat.** 1984. Deletions in the genomes of pseudorabies virus vaccine strains and existence of four isomers of the genomes. J. Virol. **49**: 970–979.
12. **Lomniczi, B., A. Gielkens, I. Csobai, and T. Ben-Porat.** 1987. Evolution of pseudorabies virions containing genomes with an invertible long component after repeated passage in chicken embryo fibroblasts. J. Virol. **61**:1772–1780.
13. **Lomniczi, B., A. S. Kaplan, and T. Ben-Porat.** 1987. Multiple defects in the genome of pseudorabies virus can affect virulence without detectably affecting replication in cell culture. Virology **161**:181–189.
14. **Ludwig, H., B. Heppner, and S. Herrmann.** 1982. The genomes of different field isolates of Aujesky's disease virus. Curr. Top. Vet. Med. Anim. Sci. **17**:15–22.
15. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
16. **Platt, K. B., C. J. Mare, and P. N. Hinz.** 1979. Differentiation of vaccine strains and field isolates of pseudorabies (Aujeszky's disease) virus: thermal sensitivity and rabbit virulence markers. Arch. Virol. **60**:13–23.
17. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
18. **Roizman, B.** 1982. The family Herpesviridae: general description, taxonomy, and classification, p. 1–23. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
19. **Sheldrick, P., and N. Berthelot.** 1974. Inverted repetitions in the chromosomes of herpes simplex virus. Cold Spring Harbor Symp. Quant. Biol. **39**:667–678.
20. **Southern, E. M.** 1975. Deletion of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
21. **Stevely, W. S.** 1977. Inverted repetitions in the chromosomes of pseudorabies virus. J. Virol. **22**:232–234.
22. **Wu, C. A., L. Harper, and T. Ben-Porat.** 1986. *cis* Functions involved in replication and cleavage-encapsidation of pseudorabies virus. J. Virol. **59**:318–327.