Genome Structure of Cottontail Rabbit Herpesvirus

JOSÉ CEBRIAN,†* NICOLE BERTHELOT, AND MADELEINE LAITHIER

Institut de Recherches Scientifiques sur le Cancer, E.R. 272, Centre National de la Recherche Scientifique, 94002-Villejuif Cédex, France

Received 9 June 1988/Accepted 10 October 1988

The genome structure of a herpesvirus isolated from primary cultures of kidney cells from the cottontail rabbit Sylvilagus floridanus was elucidated by using electron microscopy and restriction enzyme analysis. The genome, which was about 150 kilobase pairs long and which had an average G+C composition of 45%, consisted of two regions with unique base sequences (54 and 47 kilobase pairs) enclosed by reiterations of a 925-base-pair sequence with a variable copy number. The internal repeats were in opposite polarity with respect to the terminal repeats, and both unique regions underwent inversion. The nucleotide sequence of the repeat unit was determined, and virion DNA termini were precisely localized within this sequence. Elements showing homology with the cleavage-packaging signals common to other herpesviruses were detected. The data indicate that this virus is different from the previously described herpesvirus sylvilagus.

MATERIALS AND METHODS

Virus growth and purification. The cottontail rabbit herpesvirus (CTHV) was isolated from primary cultures of kidney cells from a cottontail rabbit (S. floridanus) by G. Orth (Institut Pasteur, Paris, France). Virus was grown in the rabbit kidney cell line RK13 (American Type Culture Collection, Rockville, Md.) and was cultured in Eagle minimum essential medium supplemented with 5% calf serum. Confluent cell monolayers were infected with plaque-purified virus at a low multiplicity (0.1 PFU per cell); [3H]thymidine (0.02 μCi/ml) was added to the culture medium when approximately 10% of the cells showed a cytopathic effect. Virus was harvested when more than 90% of the cell layer was affected (usually 8 to 10 days after infection).

Medium and cells were collected and centrifuged at 700 × g for 10 min, and a Dounce homogenizer was used to disrupt cells in a pellet that was previously suspended in a small volume of distilled water (1 ml, for the equivalent of 5 × 10⁶ cells in the starting culture). Debris was sedimented at 3,000 × g for 10 min. Virus from the pooled supernatants of the two centrifugations was pelleted at 12,000 × g for 2 h and suspended in 2 × NPE (1 × NPE is 10 mM sodium phosphate [pH 7.5], 100 mM NaCl, 1 mM EDTA). The Dounce-homogenized suspension was layered onto 15 to 30% sucrose gradients (in 2 × NPE) in 12-ml cellulose nitrate tubes and centrifuged in a rotor (SW41; Beckman). Fractions of 0.5 ml were collected by piercing the bottom of the tubes with a large bore needle. Portions of these fractions were assayed for trichloroacetic acid-precipitable radioactivity, and DNA-containing fractions in the middle of the gradient were pooled and ethanol precipitated. The precipitate was cen-
fuged at 10,000 × g for 1 h, dissolved in 10 mM Tris hydrochloride (pH 7.5)—1 mM EDTA (TE buffer), and dialyzed against the same buffer.

Electron microscopy. Negative staining of virus preparations was performed by the phosphotungstate procedure (6). Spreading of native, single-stranded, and partially denatured DNA was done by the technique published by Kleinschmidt (27) and as modified by Davis et al. (11).

Prehybridizations and hybridizations were performed in formamide-containing buffers by standard procedures (31), and filters were washed in 0.25× SSC–50% formamide at 37°C.

RESULTS

Preliminary characterization of CTHV. A primary culture of kidney cells from a cottontail rabbit (S. floridanus) proved to be infected by a virus possessing the structural characteristics of a herpesvirus (data not shown). We refer to this virus as CTHV.

The purified viral DNA examined by electron microscopy contained double-stranded molecules with a mean contour length of 15.07 ± 0.97 PM2 units (148.4 ± 9.6 kb; Table 1). The buoyant density of intact viral DNA in CsCl was found to be 1.698 g/cm³, corresponding to a mean G+C content of 45%. Upon shearing of DNA, a band at a density of 1.715 g/cm³ (61% G+C) appeared, while the main band at 1.698 g/cm³ broadened toward the lower densities (data not shown), suggesting an uneven distribution of G+C in the CTHV genome.

Observation of self-annealed single-stranded DNA. When allowed to renature under conditions that were favorable for intramolecular events, denatured DNA self-annealed to produce forms which appeared in EM as two single-stranded circles linked through a duplex region. These foldback molecules are diagnostic of the presence of inverted repeat sequences in the genome of CTHV, as has been found previously in a number of herpesviruses.

The two circles were of unequal size, and average diameters of the smaller and larger circles were equivalent to about 50 and 65 kb, respectively (Table 1). The duplex joining them bore a single-stranded tail in many (10 of 16 of those measured) of the molecules examined. Sizes of both duplexes and tails were different on the different molecules (Fig. 1), with their lengths varying from 1.5 to 23 kb for the tails and 1.0 to 19 kb for the duplexes. Size heterogeneity in these regions was not reflected in the total length of the molecules, which, in the limits of experimental error, was found to be constant and identical to the value measured with native DNA.

The temperature of 45°C in 50% formamide required to obtain a high number of double-looped molecules could
indicate that CTHV repeats are G+C rich and possibly are poor in DNA sequence complexity. By comparison, in our hands HSV-1 inverted repeats (70% G+C) annealed at the highest rate when DNA was incubated at 37°C in 50% formamide (39). This observation and the behavior of sheared CTHV DNA in CsCl gradients suggests that there is a large difference in the G+C content between the repetitions and the rest of the genome, and led us to choose partial denaturation as an appropriate technique for studying the CTHV genome further.

Partial denaturation: linear organization of the genome. Following incubation at 56°C in the presence of 12% formaldehyde, molecules of CTHV DNA were highly denatured and two almost completely melted regions enclosed between unmelting segments were observed (Fig. 2). Both denatured regions had constant sizes, one (UL) was equivalent to about 47 kb and the other (US) was equivalent to about 54 kb. Duplex portions at the extremities and in the median part of the molecules had extremely variable lengths, with the range of variation being roughly the same for all three regions. The sum of the lengths of the G+C-rich (unmelting) regions of the individual molecules was nearly constant, and therefore so was the total length of the molecules. The data presented thus far demonstrate that the CTHV genome has its ends and median part occupied by three G+C-rich regions with variable sizes that bracket and separate two A+T-rich regions. Furthermore, terminal portions with a high G+C content contain sequences that are present in an inverted orientation in the central region of the molecules. The discrepancies between the lengths of UL and US measured in these experiments and the size of the large and small circles in the foldback molecules are easily explained by the possibility that the circles contain a variable number of repeats. Therefore, 54 and 47 kb are likely to be more accurate estimates for the sizes of UL and US, respectively.

Partial denaturation: inversion of the unique sequences. Under milder conditions (52°C and 12% formaldehyde), which gave 38% denaturation in the unique sequences, a characteristic distribution of loops made it possible to establish a physical map of both regions with a low G+C content. The total length of regions with a high G+C content was identical to the value obtained at the higher temperature. The distribution of loops showed that regions with a low G+C content were devoid of extensively repeated sequences. To determine this distribution, molecules were oriented by aligning the two unique sequences, and then histograms representing the distribution of melted regions were drawn. The histograms obtained in this way showed a symmetric, eroded pattern (Fig. 3A and C).

In other herpesvirus genomes, segments surrounded by inverted repeats are able to invert freely. One could therefore anticipate from the data presented thus far that this would also be the case for both UL and US in the CTHV genome. Visual inspection of the molecules revealed that typical patterns of denaturation loops could be aligned by turning over the unique regions in some of the molecules. When aligned in this way, both partially denatured UL and US gave rise to asymmetrical, uneven histograms showing typical regions with high and low G+C contents (Fig. 3B and D). Regarding the relative orientations of UL and US, whole molecules could be arranged into four classes (UL-UL, UL-US, US-US, and US-US) containing 7, 7, 5, and 10 of the 29 molecules, respectively, used to derive the map. These values are statistically close to a theoretical distribution of 25% in each class.

Restriction endonuclease analyses. Except for the presence of inverted repeats, the observations made on CTHV DNA are highly reminiscent of the genome organization of herpesvirus saimiri. As described above, herpesvirus saimiri is composed of a long A+T-rich region surrounded by reiterated G+C-rich terminal sequences with heterogeneous lengths.

Restriction endonuclease analyses gave direct evidence for the presence of repetitive DNA sequences in the CTHV genome (Fig. 4). Digestion of CTHV DNA with KpnI, PvuII, or ScaI produced, in addition to molar bands, one overrepresented fragment of about 950 bp in each case; two such multimolar bands were detected in BglII (850 and 100 bp; the latter fragment is not visible in the gel shown in Fig. 4) or Smal (600 and 280 bp) digests. From the observation that KpnI, PvuII, and ScaI cut repetitive DNA into a unique class of multimolar fragments, it can be inferred that repeat units are tandemly linked in CTHV DNA molecules: cleavage of units arranged with alternate polarities in most instances would produce several multimolar fragments. Moreover, the three enzymes generated fragments with identical

<table>
<thead>
<tr>
<th>DNA</th>
<th>Contour length(a)</th>
<th>Approx. size (kb)(b)</th>
<th>% SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native DNA</td>
<td>15.07 ± 0.97 (PM2)</td>
<td>148.4 ± 9.6</td>
<td>6.4</td>
<td>39</td>
</tr>
<tr>
<td>Self-annealed single strands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small circle</td>
<td>9.87 ± 0.96 (αX)</td>
<td>53.1 ± 5.1</td>
<td>9.7</td>
<td>16</td>
</tr>
<tr>
<td>Large circle</td>
<td>12.45 ± 1.50 (αX)</td>
<td>67.0 ± 8.1</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Duplex junction</td>
<td>0.78 ± 0.51 (PM2)</td>
<td>7.7 ± 5.0</td>
<td>65.4</td>
<td></td>
</tr>
<tr>
<td>Single-stranded tail</td>
<td>1.88 ± 1.52 (αX)</td>
<td>10.1 ± 8.2</td>
<td>80.8</td>
<td>10/16</td>
</tr>
<tr>
<td>Total length</td>
<td>140.6 ± 9.1</td>
<td></td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Partially denatured DNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left duplex</td>
<td>1.27 ± 1.02 (PM2)</td>
<td>12.5 ± 10.0</td>
<td>80.3</td>
<td>15</td>
</tr>
<tr>
<td>Long denatured region</td>
<td>5.51 ± 0.22 (PM2)</td>
<td>54.3 ± 2.2</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Central duplex</td>
<td>1.68 ± 1.14 (PM2)</td>
<td>16.6 ± 11.2</td>
<td>67.8</td>
<td></td>
</tr>
<tr>
<td>Short denatured region</td>
<td>4.78 ± 0.14 (PM2)</td>
<td>47.1 ± 1.4</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Right duplex</td>
<td>1.85 ± 1.52 (PM2)</td>
<td>18.2 ± 14.9</td>
<td>82.1</td>
<td></td>
</tr>
<tr>
<td>Sum of duplex regions</td>
<td>4.80 ± 1.18 (PM2)</td>
<td>47.3 ± 11.6</td>
<td>24.5</td>
<td></td>
</tr>
<tr>
<td>Total length</td>
<td>15.09 ± 1.39 (PM2)</td>
<td>148.6 ± 13.7</td>
<td>9.2</td>
<td></td>
</tr>
</tbody>
</table>

\(a\) Length measurements are expressed relative to the length of αX or PM2 DNA (indicated in parentheses).

\(b\) Approximate sizes converted to kilobase pairs were calculated assuming that 1 unit for αX DNA = 5,386 bases and 1 unit for PM2 DNA = 9.85 kb.
sizes (950 bp), which is therefore very likely to represent the length of the entire repeat unit; the sum of the lengths of BglII and Smal multimolar fragments is consistent with this value. The relative amount of DNA contained in these bands, as estimated from densitometric scanings, corresponds to about 50 copies of the repeat unit per DNA molecule, which in total is 45 to 50 kb of the 150 kb of the genome.

Because of their variable sizes, the terminal and central regions observed by electron microscopy experiments were expected to be contained within submolar restriction fragments. In agreement with this expectation, profiles of CTHV DNA digested with BamHI, BglII, EcoRI, or HindIII showed a smear which resolved into discrete submolar bands down the gel. No multimolar fragments were observed in the digest, confirming that these enzymes do not cleave within the repeat unit. Repetitive DNA was shown to be contained in the ladders of bands by hybridization to a cloned PvuII repeat (Fig. 4).

We conclude from these results and the electron microscopy data that the terminal and central regions in CTHV DNA molecules consist of tandem repetitions of a 950-bp sequence. Internal repeats are in opposite orientation with respect to the ones at the extremities. A variation in the sizes of these regions is easily explained by a variation in the number of repeat units they contain. The overall structure that we propose for the CTHV genome is shown schematically in Fig. 5.

Nucleotide sequence of the repeat unit. The PvuII repeat unit fragment was isolated from virion-extracted DNA (see above). The entire fragment was cloned into the Smal site of plasmid pUC13, and mixtures of subfragments that were produced by cleavage with BglII or Smal were cloned into the Smal site of bacteriophage M13mp19. M13 clones were used to determine the complete nucleotide sequence of the repeat by the dideoxynucleotide procedure, as described above.

The sequence (Fig. 6) confirmed the data presented above. The repeat unit has exactly 925 bp, a value close to the apparent molecular size calculated from agarose gel migrations. It contains unique recognition sites for KpnI and SacI; two sites for BglII, giving rise to fragments of 828 and 96 bp; and three Smal sites, producing 605-, 279-, and 41-bp
fragments (the 41-bp fragment was undetected in our gels). The high G+C content predicted from partial denaturation experiments was also confirmed (63% G+C). A salient feature of the sequence is the presence of a perfect direct duplication of 40 bp; these short repeats were strictly contiguous in the repetitive DNA. The repeat sequence does not contain any long open reading frame or identifiable transcription signal. Despite similarities in the genome organization of both viruses, there appears to be no sequence homology between CTHV and herpesvirus saimiri repeats.

Localization of genomic termini. In profiles of restriction digests with enzymes cutting within the repeat unit (such as PvuII, BglII, KpnI, SacI, or Smal), the cloned repeat hybridized not only to multimolar bands but it also hybridized to a set of molar bands (Fig. 4). According to the proposed genome structure, the corresponding fragments are thought to arise, on the one hand, from the four junctions between repetitive DNA and unique sequences (U1 and U3) and, on the other hand, from the extremities of the molecules. These latter fragments are segments of the multimolar fragments, and determination of their lengths should allow for the localization of the genomic termini with respect to the repeat unit sequence.

The approximate sizes of these fragments, as calculated from their migration in agarose gels, were 490 and 340 bp with BglII, 850 bp with KpnI, 800 bp with PvuII, 580 and 350 bp with SacI, and 460 and 140 bp with Smal. Not all the corresponding bands can be seen on the gel in Fig. 4, in which KpnI and PvuII bands were mingled with the multimolar bands and the Smal 140-bp fragment was excluded, but they could readily be observed in 1.5% agarose gels (data not shown). The lengths measured from the respective sites could only be made to coincide in a disposition locating the genomic termini within an interval of 20 bp in the region which contains the 40-bp repeats.

Precise localization of termini was achieved by using high-resolution, thin polyacrylamide gels (Fig. 7). CTHV DNA digested with Smal or BglI was electrophoresed in sequence-type denaturing gels together with products of sequencing reactions, which were used as length markers. After transfer to a nylon membrane (8), DNA was hybridized to radioactive probes corresponding to either strand of the

FIG. 2. CTHV DNA partially denatured at 56°C. Transitions between the two extensive denaturation loops and the three unmelted segments are indicated by arrows on the photographs and the corresponding line drawings. Note the different sizes of duplex regions in the two molecules. Magnification, ×6,000.
528 CEBRIAN ET AL.

PvuII-SmaI fragment (coordinates 348 to 605 in Fig. 6) cloned into M13mp19. This fragment overlaps the region where the termini were found to be located by gross examination.

The lengths of the fragments detected with the lower-strand probe were 141 nucleotides for the SmaI digest and 119 and 166 nucleotides for the BclI digest (Fig. 7B). Beginning at their respective cleavage sites (positions 605 and 630, respectively), both SmaI (141 nucleotides) and BclI (166 nucleotides) fragments ended at position 465, thus defining the left extremity. In the same way, the BclI (119 nucleotides) fragment from the cut site at nucleotides 345 to 464 allowed us to map the right terminus to this latter position. Although cutting by SmaI generated blunt ends, the SmaI left end fragment cut by the upper-strand probe was found to be one nucleotide shorter than its counterpart on the lower strand (Fig. 7C). Consistently, taking into account the four-base 5' overhang left by BclI cleavage, the sizes of the upper-strand fragments (161 and 124 nucleotides) indicate that the complementary nucleotide missing at the right end is present at the left terminus. From these data CTHV DNA extremities appear to possess 3' single-base extensions, as has been described for some other herpesviruses (12, 34, 42).

DISCUSSION

In this report we described the genomic structure of a herpesvirus (provisionally called CTHV) infecting a cotton-tail rabbit, S. floridanus. This virus was isolated from a primary culture of kidney cells by G. Orth (Institut Pasteur) in the course of a search for papillomaviruses. Its genome consists of a double-stranded DNA molecule of 148 kbp containing numerous reiterations of the same set of sequences at both termini in the same orientation and internally in an inverted form. The repeats bracket two regions with unique base sequences that appear to invert freely, giving rise to four isomeric forms that are equally represented in the population of viral DNA molecules. A second source of variation in the genome structure is the length

FIG. 3. Partial denaturation maps of CTHV DNA. Histograms showing the distribution of denatured segments along the two unique regions. (A) and C) Maps of the long (UL) and short (US) regions from raw data. (B) Map of UL after inverting the direction of 14 of 29 molecules. (D) Map of US after inverting the direction of 12 of 29 molecules. The four subpopulations defined by the relative orientation of UL and US contain 7, 7, 5, and 10 of the 29 molecules.

FIG. 4. Restriction enzyme analysis of CTHV DNA. CTHV DNA was digested with the following enzymes: BamHI, BglII, EcoRI, HindIII, BglII, KpnI, PvuII, SacI, and SmaI. (A) Photograph of the ethidium bromide-stained gel; a lambda DNA-BseII digest was added as a marker. (B) Autoradiograph of the transferred gel following hybridization with the repeat unit PvuII fragment.
variability of the three repeated regions, which randomly share a constant total number of repeat units.

The CTHV genome structure appears to be an original variation on themes already encountered among herpesviruses; it can be formally derived from the genome of herpesvirus saimiri-like viruses (15), to which it is more structurally related, by the mere insertion of copies of the terminal tandem repeats into the region with a unique sequence. Accumulated data on herpesvirus genome structure and replication readily offer a basis to interpret the peculiar sequence arrangement found in CTHV. Replication of herpesvirus DNA produces concatemeric forms that are matured to unit-length genomes and encapsidated through a pathway that is supposed to involve both site-specific cleavage and "headfull" length measurement (2, 3, 7, 25, 30, 32). Such a mechanism satisfactorily accounts for features of sequence organization in various viruses (1, 5, 16, 17, 26) and abundant experimental data on normal or defective viral DNA replication (13, 29, 41, 43, 45, 46). In the case of CTHV (and herpesvirus saimiri), this mechanism explains the variable number of units in each of the repetitive regions while it maintains a constant genome length; the cleavage and encapsidation process that is initiated within a randomly selected repeat unit would package the number of reiterations between the two unique regions found in that portion of the concatemer and terminate, accommodating the number of repeats necessary to complete a genome length. The actual variation in the number of repeats in a given region of the molecule supposedly stems from recombinational events, which are also expected to promote the inversion of the unique sequences. In this connection, since recombination between direct repeats would result in deletions and genome instability, it is intuitively expected that central repeats should be in opposite polarity relative to their terminal counterparts.

A total of 10 to 15% of the virus particles produced during herpesvirus saimiri replication contain defective genomes consisting of repetitive DNA only (H genomes [5]). In the course of our experiments on CTHV, we were not able to detect such molecules, which still could exist if one assumes that recombination can accumulate enough repeat units locally to satisfy the length requirement for encapsidation. A plausible explanation would be that the frequency of this phenomenon is lowered by the distribution of the repeats into three regions in CTHV DNA instead of two in herpesvirus saimiri (more accurately, two regions instead of one in the concatemers).

Experiments on engineered genomes of HSV-1 (43) or constructed defective vectors (13) have revealed that two signals define the cutting sites within the a sequence (the terminal repeat of HSV-1), with the actual position of cleavage on each strand being independently dictated by the distance to one of these signals. Furthermore, comparison of published terminal or concatemeric junction sequences indicates that among herpesviruses these signals are strikingly conserved within otherwise totally diverged sequences (a compilation of these sequences can be found in references 13 and 18). The first signal, named pac-1 by Deiss et al. (13), is located about 40 nucleotides from one end and is composed of a T-rich element encompassed by monotonous stretches of C and G residues. The second signal, pac-2, which is about 30 nucleotides from the other terminus, contains the so-called CG motif (CGCCGCG in HSV-1) followed by a run of T residues. Sequences presenting homology with the pac-1 and pac-2 signals can be recognized at proper distances from the cleavage site that we determined. The sequence G<sub>6</sub>TGTGTTC (nucleotides 513 to 528 in Fig. 6),
Herpesvirus sylvilagus elicits lymphoproliferative disorders in cottontail rabbits, presenting similarities with the pathology of Epstein-Barr virus infection in humans, and has therefore been grouped with the other lymphocyte-associated viruses in the subfamily *Gammaherpesvirinae*. The question arises as to whether CTHV is a second isolate of herpesvirus sylvilagus or whether it is a different virus. We have no data on the biology of CTHV in its natural host, and no pathology or clinical signs could be correlated with the presence of the virus in the sacrificed animal. However, analysis of viral DNA gives a basis for comparison. CTHV and herpesvirus sylvilagus DNA molecules differ both in size and G+C content, with CTHV being about 25 kbp larger and having a lower G+C ratio (45%) compared with the G+C ratio of 52% reported for herpesvirus sylvilagus (10). Furthermore, the restriction profiles of CTHV DNA that we have presented here are completely different from those obtained by Medveczky et al. (33) and Rouhandeh and Cohrs (36), who worked with the original isolate of Hinze (21).

Finally, as an interpretation of their partial mapping data, Rouhandeh and Cohrs (36) propose for the herpesvirus sylvilagus genome an organization similar to that of Epstein-Barr virus, which is quite distinct from the genome structure we have described here. Consequently, herpesvirus sylvilagus and CTHV should definitely be considered as two different viruses.

**ACKNOWLEDGMENTS**

We are grateful to Peter Sheldrick for continuing interest in this work and to Gerard Orth for the gift of the CTHV isolate. Thanks are due to Patricia Baldacci, Suzanne Chousterman, Chantal Cremini, Luisa Dandolo, and Michel Lacasa for advice and discussion and to Thierry Foulon for help in preparing the figures.

This work was supported by grants from the Centre National de la Recherche Scientifique (ATP "Organisation et expression du génome dans les cellules eucaryotes") and the Association pour la Recherche sur le Cancer.

**LITERATURE CITED**


