A Tandem Repeat Sequence Found in a Heterogeneous Fragment of U₇ of Herpes Simplex Virus Type 1

By HITOSHI SAITO, HIROSHI SAKAOKA, TOSHIHARU YAMASHITA AND KEI FUJINAGA.

1 Department of Oral Bacteriology, School of Dentistry, Hokkaido University, Sapporo 060 and 2 Department of Molecular Biology, Cancer Research Institute, Sapporo Medical College, Sapporo 060, Japan

(Accepted 10 October 1988)

SUMMARY

We found a tandem repeat sequence in the region (designated BS7) in which restriction fragments BamHI D and Salf B overlap each other, near the centre of the unique long sequence (UL) of the herpes simplex virus type 1 (HSV-1) strain F genome. The SmaI physical map of BS7 was constructed, and the position of a heterogeneous SmaI subfragment from HSV-1 isolates and plaque-purified clones from a single strain was defined on the map. The maximum size difference in the SmaI subfragment was estimated to be 600 bp between these isolates and 100 bp between the clones. The 0.23 kb SmaI subfragment recloned from BS7 was sequenced, and was shown to contain a tandem repeat sequence consisting of 15 units of 12 bp, 5' TTGGGGCTGGGG 3'. These results suggest that the fragment length heterogeneity in the UL of HSV-1 isolates and clones is attributable to copy number variation of the tandem repeat sequence.

The genome of herpes simplex virus type 1 (HSV-1) is a linear duplex DNA of approximately 155 kb, consisting of two regions of unique sequence (UL and U₇) each flanked by two pairs of repeat regions (TR₇ and IR₇, and IR₆ and TR₆; see Fig. 1). Restriction endonuclease analyses have revealed variability in a number of HSV-1 and HSV-2 strains, representing not only gain or loss of cleavage sites, but also heterogeneity in the size of particular fragments (Wagner & Summers, 1978; Lonsdale et al., 1980; Davison & Wilkie, 1981; Umene et al., 1984a; Sakaoka et al., 1987a, b). These heterogeneous fragments are known to be derived from the regions spanning the unique sequence–repeat sequence junctions and repeated regions, among epidemiologically unrelated strains and plaque-purified clones (Wagner & Summers, 1978; Lonsdale et al., 1980; Davison & Wilkie, 1981; Varmuza & Smiley, 1984). Furthermore, it has been demonstrated that fragments located in the unique sequence, e.g. BamHI Z at the right end of U₇ (Umene et al., 1984b) and EcoRI L in the middle of U₇ (Peden et al., 1982) are also variable in HSV-1 strains. In these examples the fragment size variability observed in the repeated regions and U₇ was attributed to variation in the copy number of tandem repeat sequences (Davison & Wilkie, 1981; Mocarski & Roizman, 1981; Umene et al., 1984b). However, the causes of fragment size variability in UL have not yet been defined.

During the course of our molecular–epidemiological studies of HSV-1 and HSV-2 strains (Sakaoka et al., 1987a, b), we have also observed size polymorphism of particular fragments located near the centre of UL, i.e. in BamHI D, Salf B, PstI N of HSV-1 strains (H. Sakaoka et al., unpublished data). Of particular interest to us was whether the heterogeneity of these fragments in HSV-1 isolates could be used as a criterion for strain differentiation. By studying this size polymorphism, we have tried to determine what causes the size variability. In this paper, we report the location of a heterogeneous site on the physical map of the genome and a novel G+C-rich tandem repeat sequence found in this region.

0000-8486 © 1989 SGM
Short communication

Fig. 1. An outline of genome structure and cleavage maps of HSV-1 DNA. Heterogeneous fragments BamHI D, SalI B and PvuII N in HSV-1 isolates are indicated as heavy lines. The fragment nomenclature for BamHI and SalI is based on those of Locker & Frenkel (1979) and Post et al. (1980). The PvuII map is based on that of strain 17 published by Sanders et al. (1982). The lower part shows the region in which BamHI D and SalI B of strain F genomes overlap each other (BS7). Heterogeneous subfragments digested with BglI, Ball and Smal are also shown as heavy lines. The bottom line represents the length of BS7.

Strains F (Locker & Frenkel, 1979; Post et al., 1980; Mocarski & Roizman, 1981) and 17 (Glasgow) (Davison & Wilkie, 1981; McGeoch et al., 1985) of HSV-1 were used as standard laboratory strains. Five HSV-1 isolates from approximately 200 epidemiologically unrelated isolates (Sakaoka et al., 1987a) and three plaque-purified progeny stocks originating from a single isolate (nHS) were chosen for comparison of the fragment size. For plaque purification, confluent monolayers of Vero cells (Flow Laboratories) in 60 mm Petri dishes were infected with $10^2$ p.f.u. of the virus stock and after adsorption nutrient medium with 0.9% agar was added. Six days later, each plaque was aspirated into a Pasteur capillary and transferred to Vero cell monolayers. Plaque purification was repeated at least three times for preparation of virus clones. Confluent cultures of Vero cells in 75 cm² tissue culture flasks (Corning) were infected with each isolate or virus clone at an m.o.i. of 5 to 20 p.f.u. per cell and incubated at 36 °C for 18 to 24 h. When the cells showed maximum cytopathic changes, they were harvested for the preparation of infected cell DNA.

Preparation of DNA from infected cells and agarose gel electrophoresis were carried out essentially as described previously (Sakaoka et al., 1987a, b). Restriction endonucleases BamHI, SalI, PvuII, BglI, Ball and Smal were purchased from Takara Shuzo (Kyoto, Japan). The BamHI and SalI maps of strain F and the PvuII map of strain 17 were those of Locker & Frenkel (1979), Post et al. (1980) and Sanders et al. (1982), respectively. In order to find a heterogeneous subfragment, Southern hybridization was carried out as described (Fujinaga et al., 1979). Base sequences were determined by the dideoxynucleotide chain termination method (Sanger et al., 1977; Tabor & Richardson, 1987), using [α-32P]dATP (Amersham) and a sequencing kit (USB). Sequencing gels containing 8% and 12.5% polyacrylamide–urea (Fuji Film) were used to resolve the products.

A 7 kb DNA molecule of strain F equivalent to the overlap between BamHI D (9.6 kb) and SalI B (9.6 kb) was cloned in plasmid pUC18 (Takara Shuzo) and designated BS7. This was prepared from a cloned BamHI D fragment derived from strain F and obtained from B.
Fig. 2. Autoradiograph showing a comparison of heterogeneous Smal subfragments in HSV-1 isolates and clones. After 1 µg of each infected cell DNA had been digested with Smal, the products were electrophoretically separated and immobilized on a nitrocellulose membrane. Southern blot hybridization was carried out using 32P-labelled subfragment 0-23Sm DNA as a probe. The probe was labelled with [α-32P]dATP and [α-32P]dCTP using a random primer DNA labelling kit (Boehringer Mannheim) as described by Feinberg & Vogelstein (1983). Lanes 1 to 5 represent DNAs of five different HSV-1 isolates, A-2, A-13, A-3, A-7 and A-6, and lanes 6 to 8 indicate DNAs of three HSV-1 plaque-purified clones, nHS-1, nHS-4 and nHS-3, respectively. Size markers (M) are given in kb at the left.

Roizman (Post et al., 1980; Sakaoka et al., 1987a). We constructed the Smal physical map of BS7 DNA for further analysis of the heterogeneous region. The BamHI terminus of BS7 recloned in pUC18 was generated from the recombinant by BamHI digestion and end-labelled using [γ-32P]ATP (Amersham) and polynucleotide kinase as described by Maxam & Gilbert (1980). Thereafter, it was digested with Sall and cloned DNA was electrophoretically separated from pUC18. The Sall terminus was also labelled in the same way. These two kinds of end-labelled BS7 were partially digested with Smal, and were separated by electrophoresis in both 0.8% and 1.6% agarose gels. On autoradiographs, the sizes of end-labelled fragments were measured and the distances from each terminus to Smal cleavage sites were determined. The Smal physical map of BS7 was thus constructed and is shown in Fig. 1. The heterogeneous region was in the third Smal subfragment from the left-hand end of BS7, and was located between 1.40 and 1.63 kb from the BamHI site of BS7 (Fig. 1). Another 11 Smal subfragments were homogeneous in size among the HSV-1 isolates and clones analysed.

For determination of the extent of variation and the base sequence in this subfragment, the 0.23 kb Smal subfragment was further subcloned into Smal-restricted pUC118 (Takara Shuzo; J. Vieira & J. Messing, personal communication) and was designated 0-23Sm. DNAs from Vero cell cultures singly infected with five HSV-1 isolates and three HSV-1 plaque-purified clones from a single isolate (nHS) were digested with Smal, and electrophoretically separated on an agarose gel. Southern blot hybridization was performed using 32P-labelled 0-23Sm as a probe. Fig. 2 shows the heterogeneity in size of the fragment that hybridized. The maximum size difference of this labelled fragment was estimated to be 600 bp between strains, and 100 bp between clones of strain nHS.
Fig. 3. Autoradiographs showing the nucleotide sequence of heterogeneous SmaI subfragment 0-23Sm of strain F. The chemically degraded products of 0-23Sm were electrophoretically separated in 8% polyacrylamide-urea gels. The dried gel was exposed to Fuji XR film overnight. One copy of the tandem repeat sequence is indicated at the left of the profile.

To ascertain whether there was a reiterated sequence in the heterogeneous SmaI subfragment, the 0-23Sm subclone was sequenced by the dideoxynucleotide chain termination method. The whole sequence of 0-23Sm is shown in Fig. 3 and 4. It was found that the subfragment contained a sequence of 180 bp consisting of 15 tandem repeat units of 12 bp, 5' TTGGGGCTGGGG 3'.

As shown in the above results, we identified a SmaI restriction subfragment that varies in size among different HSV-1 isolates and particularly among different HSV-1 plaque-purified clones from a single virus population. The variable region was in the SmaI subfragment located between 1-40 and 1-63 kb from the left-hand end of BS7 (i.e. the BamHI site), where BamHI D and SfiI B overlap, at 0.47 map units (m.u.) near the centre of the U_ of strain F. In addition,
these experiments show that the 0.23Sm \textit{SmaI} subfragment of strain F contains 15 tandem reiteration units of a short sequence consisting of 12 bp (5' TTGGGGCTGGGG 3'), and has a relatively high G + C content of 75%.

Because of this high G + C content, we think it likely that the sequence is a feature similar to the following tandem repeat sequences observed in other regions of the HSV-1 genome. For example, Davison & Wilkie (1981) showed that there were three regions of reiteration in the L-S joint region of HSV-1 strain 17. The three tandemly repeated units of 12, 16 and 17 bases contain a predominance of C residues. Of these, reiteration I contained 18 copies of 5' CCGCTCCTCCCC 3', providing an explanation for the observed size variability in the a and c sequences of different plaque isolates of HSV-1 strain 17. Mocarski & Roizman (1981) also demonstrated that direct repeats 2 (DR 2) and 6 (DR 6) of strain F contain 19 copies of 12 bp tandem repeats of 5' CGCTCCTCCCCC 3' and 16 bp of 5' GCCAGGCCCTCCCCA 3', respectively, in the L-S joint region.

On the other hand, it was reported by Umene \textit{et al.} (1984b) that the \textit{BamHI} Z fragment from the U5 of strain Patton contained a unit of identical sequence but variable copy number (consisting of a 15 bp tandem repeat of sequence 5' CCACTCCCCACCCAC 3'; they showed the complementary strand) in single plaque isolates. Those authors proposed a slippage-repair model for the generation of variable copy numbers of tandem repeats during genome replication. In the same region of U5 that Umene \textit{et al.} (1984b) identified as being variable, McGeoch \textit{et al.} (1985) also found various copy numbers of a different 15 bp sequence (5' TGGGTGGGTGGGGAG 3') in HSV-1 strain 17. It is clear that copy numbers of tandem repeats differ among strains as well as among clones from a single strain, and that their difference could account for the appearance of fragment size variability.

This report shows that a tandemly reiterated sequence exists in the U1 of HSV-1 strains. We do not yet have evidence as to whether this reiteration sequence is included within a gene coding region. However, Batterson \textit{et al.} (1983) reported that a late (\(\gamma\)) gene coding for a virion component was located in the \textit{BamHI} D fragment (0.455 to 0.515 m.u. in strain F), and Zhang & Wagner (1987) reported that a 9.0 kb \(\gamma\) gene was located at 0.47 to 0.53 m.u. (strain KOS). From these reports, we consider that the reiteration sequence might be included within the coding region of a gene.

The maximum size difference was estimated to be 600 bp (50 tandem repeats) in different isolates of HSV-1 and 100 bp (eight tandem repeats) in different clones of HSV-1 strain nHS. The size polymorphism of fragments containing this reiteration region cannot be used for strain differentiation, since there is variation even among clones. Therefore restriction endonuclease sites that are gained or lost (Sakaoka \textit{et al.}, 1987a, b) should be carefully chosen for comparative analysis and for defining genetic relationships among HSV isolates.

After this paper was submitted, we learned that the whole sequence of the U1 region of HSV-1 strain 17 had been determined and that a tandem reiteration in the same region of strain 17 had been found (McGeoch \textit{et al.}, 1988). Our work was performed independently and showed that the reiteration of strain F had a different sequence and copy number from that of strain 17. Thus the conclusion on strain variability reached in this paper remains valid.
Short communication

We would like to thank Dr Bernard Roizman for providing clones of recombinant plasmids containing HSV-1 BamHI fragments. This work was partly supported by grants for research from the Ministry of Education, Science and Culture of Japan.

REFERENCES


(Received 6 June 1988)