Two Distant Clusters of Partially Homologous Small Repeats of Epstein-Barr Virus Are Transcribed upon Induction of an Abortive or Lytic Cycle of the Virus

U. K. Freeze,* G. Laux, J. Hudewentz, E. Schwarz, and G. W. Bornkamm

Institut für Virologie, Zentrum Hygiene der Universität Freiburg, D-7800 Freiburg, Federal Republic of Germany

Received 21 December 1982/Accepted 10 May 1983

The two regions of the Epstein-Barr virus genome carrying partially homologous clusters of short tandem repeats (DSL and DSR [duplicated sequences, left and right, respectively]) are transcribed into polyadenylated RNA upon spontaneous or chemical induction of the lytic virus cycle. In Raji, an Epstein-Barr virus genome carrying a nonproducer cell line, transcription of DSL and DSR is only observed upon induction of an abortive life cycle of the virus. In the nonproducer cell line Raji, the polyadenylated transcripts of DSL and DSR are about 2,500 and 2,700 bases, respectively, in length. Four different spontaneous Epstein-Barr virus producer lines, M-ABA, CC34-5, QIMR-WIL, and B95-8, differ in the length of their DSL and/or DSR regions by different numbers of tandem repeats. The size of the RNAs corresponds in all cases to the size of the respective cluster of repeats, indicating that a large part of each RNA species is colinearly transcribed from the entire tandem repeat arrays. Both the DSL and the DSR RNAs have the same polarity proceeding from right to left on the Epstein-Barr virus genome. DNA sequence analysis of the DSR repeat revealed that translation of the RNA would be possible in three open reading frames within the repeat cluster. Short homologies to herpes simplex virus IR-TR sequences and to immunoglobulin switch region sequences (IgH-S) are discussed.

The Epstein-Barr virus (EBV) is a human lymphotropic herpes virus that persists latently after primary infection (for a review, see references 17 and 18).

The virion DNA is a linear, double-stranded molecule of 170 to 180 kilobase pairs (kb) carrying identical repeats of about 400 base pairs (bp) at both termini (21, 37, 54). An array of 3-kb tandem repeats separates the molecule into a short unique (US) and a long unique (UL) region of about 15 and 140 kb, respectively (21, 24, 61). In latently infected cells, the viral DNA is usually carried in multiple copies as closed circular episomal DNA (43).

Within the long unique region, two regions located about 100 kb apart from each other, denoted DSL and DSR (duplicated sequence, left and right, respectively), have been shown to share sequence homologies (56). The detailed analysis of these regions revealed the presence of small tandem repeats in DSL (23) and DSR (29). Fragments containing DSL or DSR exhibit a remarkable length polymorphism in different virus isolates (8, 29), which was shown to be due to different numbers of repeats in DSR (29). Heteroduplex analysis of recombinant plasmids containing DSL and DSR, respectively, showed that the region of homology is present in the same orientation in the viral genome and consists of about 1.5 kb with partial homology and about 0.9 kb with good (perfect) homology (29). B95-8 (48) is an unusual derivative of EBV with a deletion of 12 kb, including the complete DSR region (7, 19, 56). The structural organization of the genome of M-ABA (EBV), a virus isolated from nasopharyngeal carcinoma cells (12), is schematically described in Fig. 1.

In most EBV-harboring cell lines, the majority of cells within the cell culture population maintain the virus in a latent state. Usually, only a minority of cells spontaneously enters the lytic cycle of the virus (producer lines). In these lines, the number of virus-producing cells can be efficiently increased by treatment of the cells with the tumor promoter, 12-O-tetradecanoyl-phorbol-13-acetate (TPA) (28, 73, 74).

In the EBV genome carrying cell line Raji, late viral functions are not expressed even after treatment with TPA (nonproducer). In this cell line, early viral functions are, however, efficiently induced by a variety of compounds including TPA (28, 73). The cellular and/or viral
factors that maintain the latent state or initiate and promote the lytic phase are so far undefined.

An important approach to study viral latency and the early events of the virus life cycle after induction involves the analysis of viral transcription. Previous work identified distinct regions of the viral genome which code for EBV-specific RNA in producer and nonproducer cell lines (1, 22, 30, 32, 36, 42, 52, 59, 60, 69).

Since the viral DNA has been cloned by recombinant DNA technology in several laboratories (2, 10, 13, 64; A. Polack, G. Hartl, U. K. Freese, U. Zimber, L. Gissmann, and G. W. Bornkamm, manuscript in preparation), it is now possible to refine the transcription map of EBV with greater resolution (30, 31). A first transcription map of the B95-8 strain has been recently described (30).

In this report, we describe the identification of viral transcripts coded for by the DS_L and DS_R repeats in four producer lines and in one nonproducer cell line. These transcripts are newly synthesized upon induction of Raji cells with TPA. Additionally, we present the DNA sequence of the DS_R repeat. Besides the possible protein coding function, this sequence bears an interesting structural and possible functional relationship to immunoglobulin switch region sites (IGH-S) and related sequences in *Drosophila melanogaster* and herpes simplex virus (16, 35, 49, 62).

**MATERIALS AND METHODS**

**Cells.** All cell lines were grown in RPMI 1640 medium (GIBCO Laboratories, Grand Island, N.Y.) with 2.5% fetal calf serum and 8% calf serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml. For RNA extraction, cells were diluted twice weekly 1:2 to 1:3 with fresh medium to a concentration of 0.1
\(10^7\) to \(0.2 \times 10^6\) cells per ml of culture and harvested at \(0.3 \times 10^6\) to \(0.4 \times 10^6\) cells per ml. Cells were induced with 20 ng of the tumor promoter TPA per ml for 2 days. The degree of induction was monitored by immunofluorescence with sera containing anti-VCA and anti-EA antibodies. The degree of induction varied in the different cell lines between 5 and 10%. M-ABA cells were less inducible and reached never more than 3% VCA-positive cells. The uninduced controls showed about 1 to 3% spontaneous induction for the producer lines. For the nonproducer Raji cell line, the background for spontaneous EA induction was \(\leq 0.01\%\). Viability of the cells was usually \(\geq 95\%\).

Initial cultures were obtained from M. A. Epstein (M-ABA) (12), W. Henle (Raji) (55), G. Miller (B95-8 and CC34-5) (47, 48), and J. Pope (QIMR-WIL) (53).

**RNA preparation and RNA blot analysis.** RNA was extracted from the cytoplasmic fraction of cells as described by Kumar and Lindberg (41) and Berk and Sharp (6). The protein extraction of cytoplasmic supernatants was carried out by the method of Holmes and Bonner (27). Polyadenylated RNA was purified from cytoplasmic RNA preparations essentially as described by Jelinek et al. (34).

RNAs were electrophoresed on denaturing, horizontal agarose slab gels as described by Seed (63). The RNAs were transferred to and immobilized on nitrocellulose with a pore size of 0.1 \(\mu\)m (PH 79, Schleicher & Schuell Co., Dassel, Federal Republic of Germany) according to Thomas (68). Cloned DNA fragments of M-ABA (EBV) or purified restriction fragments of the cloned DNA were labeled by nick translation with \(^{32}\)P-dCTP (40 Ci/mmol; Amersham Corp., Arlington Heights, Ill.) in a total volume of 20 to 50 \(\mu\)l (45, 57).

The specific activity of labeled probe was between 5 \(\times 10^4\) and 10\(^5\) cpm/\(\mu\)g of DNA. Hybridization, washing after hybridization, and autoradiography were carried out as described previously (8). The hybridization temperature was 45°C unless otherwise stated in the text. RNAs (28S, 5,100 bases [b]; 18S, 1,800 b) served as internal size standards (11). Additionally, a DNA marker mix was used containing linear pH79 dimer and monomer, pBR322 linear monomer, and additional digests of pBR322 with BamH1-PstI and with HinfI. The sizes of the different marker bands are either given in the figures or stated in the figure legends. The DNA marker bands were detected on the blots by hybridization of the vector part of the labeled probes.

**Digestion with restriction enzymes and gel electrophoresis.** Restriction enzymes were purchased from Boehringer Mannheim Corp., Mannheim, Federal Republic of Germany; Bethesda Research Laboratories, Bethesda, Md.; and New England Biolabs. New England Biolabs indicates that the restriction endonuclease NotI (R. Borsetti, D. Wise, and T. Atkinson, unpublished data; 58) recognizes the 10-b palindrome GCC\(^4\)GCGGCC, that the internal hexanucleotide palindromes is an XmalII site, that not all XmalIII sites are substrates for NotI, and that not all NotI sites are substrates for XmalII. Digestion of DNA with restriction endonucleases, separation of fragments by gel electrophoresis, and Southern transfer to nitrocellulose or isolation of fragments from the gel were carried out essentially as described previously (8).

**DNA sequence analysis and preparation of single-stranded hybridization probes.** Sequence analysis and preparation of \(^{32}\)P-labeled single strands for directional hybridization of RNA blots were carried out by the method of Maxam and Gilbert (46). Several different computer programs were used to exploit the sequence data (9, 40, 51, 66).

**RESULTS**

Induced transcription of DS\(_L\) and DS\(_R\) in Raji and M-ABA. The structural organization of the M-ABA (EBV) genome and the probes used for hybridization are schematically shown in Fig. 1. A detailed description of the cloned probes will be published elsewhere (A. Polack, G. Hartl, U. K. Freese, U. Zimber, L. Gissmann, and G. W. Bornkamm, manuscript in preparation).

To study the effect of TPA on the transcription of the Raji (EBV) genome, total cytoplasmic RNA (10 \(\mu\)g) from induced (Fig. 2) and uninduced Raji cells was run on a denaturing gel (63), transferred to nitrocellulose (68), and hybridized with \(^{32}\)P-labeled M-ABA viral RNA labeled by nick translation. One prominent intense broad RNA band in the range of 2,500 to 2,700 b between the 28S and the 18S endogenous marker was found to be induced among a series of other fainter bands. The relatively intense band at the bottom corresponds to the recently described virus-coded small ribonucleoprotein particle-bound RNA of high abundance (42, 59, 60). The intensity of this band did not change after induction.

In an attempt to map the abundant inducible transcript on the viral genome, five overlapping recombinant cosmid clones of M-ABA (EBV) DNA and a series of subclones derived from the two regions of overlap between cosmid clones cM401-3 (1 in Fig. 1), cM302-17 (2 in Fig. 2), cM405-15 (4 in Fig. 1), and cM302-21 (5 in Fig. 3) were hybridized to RNA blots containing cytoplasmic RNA of induced and uninduced nonproducer Raji and M-ABA producer cells. The overlap of cosmid clones 1 and 2 contains the DS\(_L\) region, and the overlap of cosmid clones 4 and 5 contains the DS\(_R\) region of the EBV genome (Fig. 1). The results of the different blot hybridizations are summarized in Table 1 and are shown in part in Fig. 2.

In M-ABA and Raji cells, only probes containing DS\(_L\) (probes a and g in Fig. 1) or DS\(_R\) (probes n and q in Fig. 1) sequences hybridized with the abundant inducible transcript (Fig. 2, Table 1). In M-ABA, this RNA is already present without TPA treatment. After short-term exposure, the broad, intense hybridization signals of M-ABA and Raji RNA are resolved into two bands of about 2,700 and 2,500 b in RNA from Raji and about 2,800 and 2,600 b in RNA from M-ABA cells (Fig. 2). The sequences flanking the DS\(_L\) region in clone g (Fig. 2A) are too short to code for an RNA of a minimum of 2,500 b. It is thus concluded that sequences of the cross-hybridiz-
FIG. 2. (A) Autoradiograms of blots containing size-fractionated cytoplasmic RNAs (10 μg per lane) of chemically induced (+) and uninduced (−) Raji (R) and M-ABA (M) cells after hybridization with \(^{32}P\)-labeled M-ABA virus DNA (M-vDNA) and with different recombinant plasmid clones containing DS\(_{L}\) (probes a and g in Fig. 1) and DS\(_{R}\) (probes n and q in Fig. 1). The two leftmost panels are short-term exposures. rRNA 28S (5,100 b) and 18S (1,800 b) are indicated at the sides of the blots by a dot. (B) Autoradiograms of blots containing polyadenylated RNA purified from induced (+) and uninduced (−) M-ABA (M) and Raji (R) cells. The \(^{32}P\)-labeled hybridization probes are described in the text and Fig. 1. Each lane contains about five times the amount of polyadenylated RNA (≈1 μg) which is contained in 10 μg of unfractionated cytoplasmic RNA. Hybridization temperature was 50°C. Molecular weights (in b) are indicated at the sides of the outer blots.

In this clone, the DS\(_{R}\) region is flanked by unique sequences of 650 bp to its left and 960 bp to its right-hand side (29). Thus, any transcript with a minimum of 2,500 b which is transcribed
from this region must contain a considerable part of the cross-hybridizing DS_R region. This implies that a transcript encoded completely or partially by DS_L or DS_R must hybridize to DNA from both regions.

The DS_R transcript is coded mainly by the short tandem repeats. The DS_R region in M-ABA (EBV) is composed of about 25 short tandem repeats which are partially homologous to the DS_L repeats and are about 900 bp adjacent to the repeats on the right-hand side with almost perfect homology to sequences in DS_L. To determine which part of the DS_R region is represented in the abundant EBV-specific RNA, we attempted to map the transcript more precisely in DS_R. For this reason, the DS_R repeat was cloned into the PstI site of pHC79 (probe t in Fig. 1). The two BglII-PstI fragments flanking the cluster of repeats were isolated from a preparative agarose gel (probes s and u in Fig. 1). The result of the hybridization for M-ABA and Raji polyadenylated RNA with these three 32P-labeled probes is shown in Fig. 2 and Table 1.

The left flanking fragment of about 700 bp and the probe of the DS_R repeat hybridized about equally well to the abundant RNAs (Fig. 2B). Since the left-flanking BglII-PstI fragment contains almost a complete DS_R repeat unit (see below), it cannot be decided whether the hybridization of this fragment to the RNA is due to the repeat alone or is additionally due to the unique flanking sequences. The fragment to the right of the repeats, however, which contains the region of perfect homology hybridized considerably less. This suggests that the abundant RNA transcribed from DS_R contains mainly sequences complementary to the DS_R repeat cluster.

DS_L and DS_R each code for individual transcripts. The probes of DS_L and DS_R visualized different patterns of RNA in M-ABA and Raji cells (Fig. 2). The probes of DS_L and DS_R hybridized to two RNA species of 2,500 and 2,700 b from Raji cells and to two species of 2,600 and 2,800 b from M-ABA cells. The DS_L probe hybridized more prominently to the transcript in the lower band and less to the RNA in the upper band of Raji and M-ABA cells. Inversely, the DS_R probe hybridized strongly to the RNA in the upper band and weakly to the RNA of the lower band. This suggested that both DS_L and DS_R might code for individual transcripts which react to different extents with both probes due to homology in their nucleotide sequence.

To test this, we looked for RNA complementary to DS_L and DS_R in B95-8 cells. B95-8 virus has a deletion of about 12 kb, which includes the complete DS_R region. Hybridization of B95-8 RNA revealed a single band of 2,400 b which reacted strongly with the DS_L probe and faintly with the probe containing DS_R (Fig. 4). Since in this case the RNA must be transcribed from
DS_R, the hybridization with the DS_R probe must be due to partial sequence homologies of the RNA with the DS_R probe. Increase in the hybridization temperature from 45 to 50°C reduced the intensity of the hybridization with the DS probe, although the hybridization of the RNAs of the other strains was not affected (Fig. 4). At 65°C the cross hybridization of the DS_R probe with the transcript from B95-8 cells was completely abolished (Fig. 4).

RNA from two additional nondefective producer lines, CC34-5 (47) and QIMR-WIL (53), were studied for comparison (Fig. 4). RNA from CC34-5 cells showed one band of about 3,300 b, which hybridized equally well to probes of DS_L and DS_R. In contrast, in QIMR-WIL RNA, the DS_L probe hybridized strongly to a band of about 2,300 b and weakly to a band of 3,300 b, whereas the DS_R probe revealed the inverse hybridization pattern. The increase of the hybridization temperature to 50°C again reduced the intensity of the hybridization of the DS_R probe to the lower band without affecting the intensity of hybridization to the upper band (Fig. 4). Further increase of the hybridization temperature to 65°C restricted the hybridization exclusively to the upper band (Fig. 4). Hybridization at 65°C with a DS_L probe restricted the annealing of the label exclusively to the lower band. Only after prolonged exposure times could the upper band also faintly be detected (data not shown).

It is concluded from these experiments that both regions DS_L and DS_R code for individual transcripts which share some sequence homologies. The failure to detect two separate bands in CC34-5 RNA presumably indicates that both transcripts from DS_L and DS_R have about the same size and are not separated in the gel.

**Polarity of the DS_L and DS_R transcripts.** To define the direction of transcription within the DS_L and the DS_R regions, we hybridized the isolated, separated, 5' end 32P-labeled single strands of the DS_L and DS_R repeats to blots containing the size-fractionated, polyadenylated RNAs of induced and uninduced M-ABA, Raji, B95-8, and QIMR-WIL cells. The hybridization was carried out at 65°C. The polarity of each single strand of DNA was defined by determination of its nucleotide sequence which allowed it to align relative to the standard EBV genome (14, 33; see Fig. 6). The result of the hybridization is shown in Fig. 5.

Only the upper strands of each of the DS_L and the DS_R repeats, with the 5' to 3' orientation from left to right, hybridize to the RNA on the blots (Fig. 5). The lower strands under the same conditions did not hybridize at all. This indicates that the DS_L and DS_R RNAs are transcribed from right to left. The upper single strand of the NotI repeat fragment of DS_L used in this experiment as the probe reacts with about equal intensity with both the DS_L and the DS_R RNAs of QIMR-WIL. From M-ABA EBV, it is known that the DS_R repeat cluster contains about 25 copies of the repeat monomer, whereas its DS_L counterpart contains only about 13 single repeat units. The complementary related target sequences for hybridization of the DS_R repeat to both RNAs are thus in twofold molar excess within the DS_R RNA as compared with the DS_L RNA. This explains the differential hybridization of this small DS_R repeat probe which lacks flanking sequences unique to the DS_L RNA which preferentially reacts with the lower band of QIMR-WIL RNA (Fig. 4 and 5).

**Size differences of the transcripts correlate with length polymorphism of DS_L and DS_R in different virus strains.** As mentioned above, the size of the RNAs transcribed from DS_L and DS_R varied markedly in different strains. Thus, the size of the DS_L RNA varied between 2,300 and 3,300 b and the size of the DS_R RNA varied between 2,700 and 3,300 b in the strains tested.
It has been shown that fragments carrying $D_{S_L}$ and $D_{S_R}$ exhibit a marked length polymorphism (7, 8, 56) and that different numbers of $D_{S_R}$ repeats account for the length polymorphism in fragments containing $D_{S_R}$ (29). Viral DNAs were digested either with BamHI alone or with BamHI and NotI, which cuts within the $D_{S_L}$ repeats (Fig. 5). The sizes (in b) of the BamHI H fragments carrying $D_{S_L}$ are: QIMR-WIL (6,100), M-ABA (5,400 plus 1,000), CC34-5 (7,000), B95-8 (6,200), and Raji (6,300). After double digestion, the sizes of the flanking fragments of the NotI repeat clusters are 2,800 bp in all producer strains and 1,700 bp in all producer strains except M-ABA. In M-ABA, an additional BamHI cleavage site within BamHI-H generates two fragments of 1,000 and 700 bp, which together have the same size as the 1,700-bp fragment of the other producer strains. In double digests, all four viral DNAs revealed fragments of the same size, whereas the size differences of the BamHI H fragments were pronounced after BamHI digestion alone. This indicates that different numbers of $D_{S_L}$ repeats account for the size variation in BamHI-H. In contrast, the $D_{S_L}$ repeats of Raji (EBV) appeared not to be cut by NotI. The recognition sequence of NotI contains two CpG dinucleotides which are the main targets for methylation in higher cells. The fact that NotI does not cut within the $D_{S_L}$ repeats of Raji (EBV) might be due to extensive methylation of Raji (EBV) DNA (38).

The lengths (in b) of the corresponding $D_{S_L}$ RNAs are: QIMR-WIL (2,300), M-ABA (2,600), CC34-5 (3,300), B95-8 (2,400), and Raji (2,500), (Fig. 4). It is obvious that the size of the transcripts correlates with the size of the corresponding DNA fragments. This indicates that the RNAs are colinearly transcribed from the complete cluster of $D_{S_L}$ and $D_{S_R}$ repeats in each strain.

**DNA sequence of the $D_{S_R}$ repeat.** To study whether the $D_{S_R}$ repeats can possibly code for a protein, we determined the nucleotide sequence...
of one DSR repeat unit. The sequence, its restriction enzyme cutting sites, and the hypothetical translation products in the three reading frames of the coding strand are shown in Fig. 6. One repeat unit is 102 bp long, contains 82% G+C, and displays three open reading frames for translation. None of the possible reading frames contains an initiation codon. Figure 6 also documents the primary sequence data of both strands of a part of the DSR repeat of M-ABA EBV. At position 58 to 63, a run of six guanosine residues occurs. Correspondingly, six cytosine residues are located in the complementary strand in the same position. Additionally,
FIG. 6. (A) Fine map of M-ABA BgIII-K (probe q in Fig. 1) which contains the DS region of the viral genome. PsI and Aval cleavage sites are indicated by vertical bars. A 178-bp part of the DS repeat cluster representing an Aval and an overlapping PsI repeat unit of 102 bp each are shown as an expanded fine map. The sequencing strategy is indicated by horizontal arrows below the map. Closed circles represent labeled 5' ends and open triangles represent labeled 3' ends of the DNA single strands which were used for sequencing. (B) Double strand of the nucleotide sequence of the 178-bp fragment of the DS repeat cluster. Restriction endonuclease cleavage sites are indicated vertically above their recognition sequences (for further details of the sequence presentation, see reference 40). A nine-bp sequence homologous to herpes simplex virus type 1 (F)-a-DR2 and switch region-like sequences are boxed. Details are given in the text and Fig. 1 and 2. The neighboring run of six consecutive guanosines, which is different from the published sequence of the DS (IR4) repeat sequence in another EBV virus strain (AG876), is underlined (14) (14 b, as described in reference 14). The DS repeat sequence is underlined by the amino acid sequences that can be formally translated in the three open reading frames of the coding strand. The orientation of the respective open reading frames is indicated by an arrowhead at the amino terminus of the peptide sequence in each line. The direction of transcription is from right to left. (C) Primary sequence data of position 97 to 40 of the coding and position 74 to 40 of the noncoding strand of the DS repeat. The six cytosines and the corresponding six guanosines are underlined and marked by black vertical arrowheads. (D) The M-ABA DSa (top line) and the AG876 IR4 (bottom line) sequences are aligned. The extra guanosine in IR4 stands out and is marked by a star. All other positions are identical.
this result was confirmed by sequencing the upper 3'-end-labeled strand from the nearby PstI site (data not shown). The recently published AG876 IR4 sequence contains seven guanosine residues at the same position (14). This extra guanosine residue is the only difference between the M-ABA and the same published AG876 sequence; all other positions of DS_R IR4 are identical. As we did not sequence the IR4 repeat of AG876 ourselves, we cannot decide whether the published sequence is different by an experimental artifact or represents a true difference among two EBV virus strains. The consequences of this difference are drastic for the primary structure of a hypothetical translation product, as it generates a frame shift of translation of plus one in every repeat unit.

**DISCUSSION**

Our data show that DS_L and DS_R each code for an individual transcript. The length polymorphism of these regions on the DNA level in different EBV strains corresponds to the size differences of the respective transcripts. Our map data and the fact that this polymorphism is due to varying numbers of repeats in DS_R and DS_L confine the transcripts to the two partially homologous tandem repeat arrays and their flanking sequences. The RNAs are polyadenylated and transcribed from right to left on the EBV genome. Comparison of the physical length of the DS_R RNA in M-ABA cells (2,800 b) and the total length of the tandem repeat array in DS_R (2,550 bp) indicate that about 98% of the colinear part of the transcript is derived from the repeat sequences if one allows for a 200-b polyadenylated tail. If the size measurements of the RNA and DNA are correct, about 50 b are left which can be derived from the flanking sequences. In DS_L the total length of the tandem repeat array (1,700 b) constitutes a relatively smaller fraction of the colinear RNA (about 70%), leaving about 700 b to be transcribed from the flanking sequences.

The results indicate that DS_L and DS_R transcription occurs either spontaneously or is chemically induced. In the nonproducer Raji cells, we were unable to detect a low level of constitutive expression of DS_L and DS_R.

In Burkitt lymphoma biopsies, abundant EBV-specific transcription has been found in the region of the viral genome which is deleted in B95-8 and contains the DS_R transcription unit (15). This could indicate that possibly DS_R-specific RNA is spontaneously induced in Burkitt tumors in vivo. The DNA structure of DS_L and DS_R suggests that within the highly conserved part of both

**Alignment of DS_R (#1) and DS_L (#2) repeat:**

| #1  | 5' CAGCTGC AGC CCGCAACCCCCGCC GAGCGGGGG CAGCGGACCCAGCGGACC C GTGCGCGCCACCC GCC 3' |
| #2  | 5' CCG TGCCCAGGCGGGCGACCCCCCGACCGCCC CAGGGCGG CGACGG CCCAGGCGGGACCC GCGCC GCGG 3' |

Maximal homology: 79.%; probability (76, 60) < 0.0001

**HSV1(F)-a-DR2 (#3)**

| # 3 5' GGGGGA GAGCGGGGG GGA GAGCGGGGG GGA GAGCGGGGG (GGA GAGCGGGGG)N 3' |

**Prevalent sequences of S_p1, S_p2b and S_p3 regions (#4)**

<table>
<thead>
<tr>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>04 Y1 5' GF FT CC AGCTG AGCTCA CA GG GAGCT TGGG G TA GTTGGATT G 3'</td>
</tr>
<tr>
<td>04 Y2b 5' GG GA CA AGCTCT AGCTCT TG GG GAGCT TGGG G A7 GTTGGATT G 3'</td>
</tr>
<tr>
<td>04 Y3 5' PG NA CC TCGTCA AGCTGCTA CA GG GAGCT 7GGG G 7P GTTGGATT G 3'</td>
</tr>
</tbody>
</table>

**S_p-like sequences of Drosophila melanogaster , sequence V (#5)**

<table>
<thead>
<tr>
<th>3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>05 5' GAGCTGGG TAGCTGGG TAGCTGGG GAGCTGGG GAGCTGGG TAGCTGGG GAGCTGGG 3'</td>
</tr>
</tbody>
</table>

*Fig. 7. Alignment of the homologous parts of the DS_R repeat (#1) and of the DS_L repeat (#2) (33). Additionally, the sequences of herpes simplex virus type 1 (F)-a-DR2 (#3), the consensus sequences of the Ig-γ-H switch regions of mice (#4), and the switch region-like sequences recently detected in D. melanogaster (#5) are shown. The related sequences are underlined.
regions, a regulatory element acts which coordinately initiates and controls transcription of the DS$_L$ and DS$_R$ RNAs (Fig. 5). A similar structural and functional organization has already been found in other eukaryotic gene systems (3, 47).

So far no data are available to define the function of DS$_L$ and DS$_R$ transcripts. It is not known whether both of them code for messenger RNAs translated into one of the already known EBV-specific early proteins. The fact that DS$_L$ and DS$_R$ transcripts are polyadenylated and can be obtained from polysomes suggests that they indeed code for proteins (30, 70). The sequence of the M-ABA DS$_R$ repeat allows for such a model, since all possible reading frames of the coding strand are in phase and open for translation. The putative gene product of the DS$_R$ repeats would consist mainly of 34 amino acids, which are repeated about 25 times. Other proteins with repetitive structures like collagen (39, 71), silkfibroin (44, 65), melting point lowering serum protein in arctic fish (72), and zein (20), the storage protein of maize, show shorter and/or fewer and less-conserved intragenic repetitions (for a review, see reference 4). Recently, it has been observed that the gene product of the Balbiani ring 2 of polytene chromosomes in *Chironomus tentans* has a comparatively conserved repeat structure (67).

The comparison of the nucleotide sequence of the DS$_R$ and the DS$_L$ repeats (33) revealed 76 bp of high homology (Fig. 7). A 9-bp sequence motive, GGAGCGGGG, of these 76 bp was also detected in the Ig-R$_I$ inversion sequence a of herpes simplex virus (16, 49) (Fig. 7). A short related sequence, GGAGCGGG, was also found in the switch site consensus sequences of the Ig-$\gamma$-H locus of the mouse (35) (Fig. 7). Similar switch region-like sequences have recently been found in *D. melanogaster*, indicating a more general importance of this type of sequence (62). One of these sequences in *D. melanogaster* contains six tandem repeats of GGAGCTGGG, the sequence shared among DS$_R$-DS$_L$, herpes simplex type 1 "a," and the IgH-S sequences of mice (Fig. 7). Analysis of the defective molecules in the P3HR-1 virus DNA population has revealed that the DS$_R$ repeat region is also a target for intramolecular recombination (M. Cho, G. W. Bornkamm, and H. zur Hausen, manuscript in preparation).

The two pseudodiploid EBV genes, DS$_L$ and DS$_R$, and their transcripts may serve as two sensitive targets for future studies of maintenance of latency and induction of the lytic cycle. Further work will be necessary to reveal their possible coding and recombinatorial functions.

After our work was submitted, Dambaugh and Kieff (14) and Jeang and Hayward (33) also presented RNA data for the transcription of the DS$_L$ (14, 33) and DS$_R$ regions (33) which are in good agreement with our own data. The AG876 IR4 sequence presented by Dambaugh and Kieff (14) differs in one position from our M-ABA DS$_R$ repeat sequence. This difference is discussed above (Fig. 6).

**ACKNOWLEDGMENTS**

We are most grateful to D. S. Hayward and K. T. Jeang for discussion of their results before publication and for their support in using the Stanford Mol Gen Computer facilities for the homology search among DS$_L$ and DS$_R$ repeat sequences. We are indebted to C. van Beveren and Inder Verma for allowing us to search the DNA sequence database at the Salk Institute for homologies of the DS$_R$ repeat to other eucaryotic sequences. We thank A. Dunn and N. Gough for many helpful discussions. The expert technical assistance of E. Kohler is gratefully appreciated. We thank H. zur Hausen for his support and for critically reading the manuscript.

This work was supported by die Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 31, Tumorentstehung, und Entwicklung.

**LITERATURE CITED**


ERRATA

Two Distant Clusters of Partially Homologous Small Repeats of Epstein-Barr Virus Are Transcribed upon Induction of an Abortive or Lytic Cycle of the Virus

U. K. FREESE. G. LAUX. J. HUDEWENTZ. E. SCHWARZ. AND G. W. BORNKAMM

Institut für Virologie, Zentrum Hygiene der Universität Freiburg, D-7800 Freiburg, Federal Republic of Germany

Volume 48, no. 3, p. 731: The present address of U. K. Freese is as follows: Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 6900 Heidelberg 1, Federal Republic of Germany.

Page 731, Abstract, lines 4 and 5: Should read “... In Raji, a nonproducer cell line carrying the Epstein-Barr virus genome, transcription . . .”

Page 731, column 1, line 17: “sequence” should read “sequences.”

Page 731, column 2, line 26: “(28, 73)” should read “(5, 28, 73).”

Page 732, legend to Fig. 1, lines 4 and 5: “... Fig. 2 through 7 and Table 1 . . .” should read “... Fig. 2 through 4, Fig. 6, and Table 1 . . . .”

Page 732, column 1, lines 8 and 9: “(1, 22, 30, 32, 36, 42, 52, 59, 60, 69, 70)” should read “(1, 22, 26, 30, 32, 36, 42, 52, 59, 60, 69, 70).”


Page 733, column 2, Results, line 10: “cytoplasmic” should read “cytoplasmic.”

Page 733, column 2, Results, line 30: “(2 in Fig. 2)” should read “(2 in Fig. 1).”

Page 733, column 2, Results, line 31: “(5 in Fig. 3)” should read “(5 in Fig. 1).”

Page 736, Table 1 heading, line 1: “different” should read “different.”

Page 737, column 1, line 8: “... (Fig. 5). The sizes (in b) . . .” should read “... (Fig. 3). The sizes (in bp) . . . .”

Page 739, legend to Fig. 6, lines 8 and 9: “... homologous to herpes simplex virus type 1 (F)-a-DR2 . . .” should read “... homologous to the direct repeat sequence DR2 within the a region of herpes simplex virus type 1 strain F [HSV-1(F)-a-DR2] . . . .”

Page 740: Fig. 7 should appear as shown below.

```
Alignment of DS_R (#1) and DS_L (#2) repeat:

#1 5' CGGTCG ACCGCACCGCCCGCC CGAGGCGGG CAGCGAACCCCGGGGG ACCGCGCGC GGGGGCGGG 3'
#2 5' CGG TGGCCCGCCCCCCGCCCCCGGGG CGAGGCGGG CAGCGCGGG CCCAGGGAGCCCCCGGG GGCGC GCCGGCCG 3'

Maximal homology: 79.2; probability (76, 60) < 0.0001

HSV1(F)-a-DR2 (#3)

#3 5' GGGGA GGAGCGGGG GGA GGAGCGGGG GGA GGAGCGGGG GGA GGAGCGGGG (GGA GGAGCGGGG)N 3'

Prevalent sequences of S_1', S_2b' and S_3' regions (#4)

#4 y_1 5' CGGTCG ACCGCAC CGAGCCTGA CA CG CGCCT GTA GGA YAY FPTGCA ATP GT 3'
#4 y_2b 5' GGGC ACCGACCTG TA CGGCGGACTA GGGCGGACTA GAT GAT GAT GT 3'
#4 y_3 5' CGGTCG ACCGACCTG TA CGGCGGACTA GGGCGGACTA GAT GAT GAT GT 3'

S_p'-like sequences of Drosophila melanogaster , sequence V (#5)

#5 5' CGGTCGCGG TGCGGCGG TGCGGCGG CGGCGGCGG CGGCGGCGG CGGCGGCGG 3'
```

Page 740, legend to Fig. 7, line 2: “... herpes simplex virus type 1 (F)-a-DR2 . . .” should read “... HSV-1(F)-a-DR2 ... . . . .”
Page 741, column 1, line 5: "(3, 47)" should read "(3, 47, 50)."
Page 741, column 1, line 50: "M. Cho" should read "M.-S. Cho."
Page 741, column 2, line 3: "Kief" should read "Kieff."
Page 741, column 2, Acknowledgments, line 3: "Mol Gen" should read "MOLGEN."
Page 742, column 1, reference 21: The article title should read as follows: DNA of Epstein-Barr virus. VI. Mapping of the internal tandem reiteration.
Page 742, column 1: Reference 25 should be deleted.

Epstein-Barr Virus-Specific DNase Activity in Nonproducer Raji Cells After Treatment with 12-O-Tetradecanoylphorbol-13-Acetate and Sodium Butyrate

T. OOKA, M. DE TURENNE, G. DE THE, AND J. DAILLIE

Laboratoire d'Epidémiologie et d'Immunovirologie des Tumeurs, Faculté de Médecine Alexis Carrel, 69372 Lyon Cedex 2, France

Volume 49, no. 2, p. 626, column 1, line 36: "4 \times 10^5 to 4 \times 10^6" should read "4 \times 10^5 to 1 \times 10^6."
Page 628, column 1, legend to Fig. 1, line 5: "footnote a" should read "footnote c."
Page 628, column 2, Literature Cited, reference 7, line 2: "injection" should read "infection."
Page 628, column 2, Literature Cited, reference 13, line 3: "6270–6271" should read "6270–6277."