

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

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 (DS_L and DS_R [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 DS_L and DS_R is only observed upon induction of an abortive life cycle of the virus. In the nonproducer cell line Raji, the polyadenylated transcripts of DS_L and DS_R 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 DS_L and/or DS_R 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 DS_L and the DS_R RNAs have the same polarity proceeding from right to left on the Epstein-Barr virus genome. DNA sequence analysis of the DS_R 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 I_R-T_R 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 (U_S) and a long unique (U_L) 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 DS_L and DS_R (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 DS_L (23) and DS_R (29). Fragments containing DS_L or DS_R exhibit a remarkable length polymorphism in different virus isolates (8, 29), which was shown to be due to different numbers of repeats in DS_R (29). Heteroduplex analysis of recombinant plasmids

containing DS_L and DS_R , 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 DS_R 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

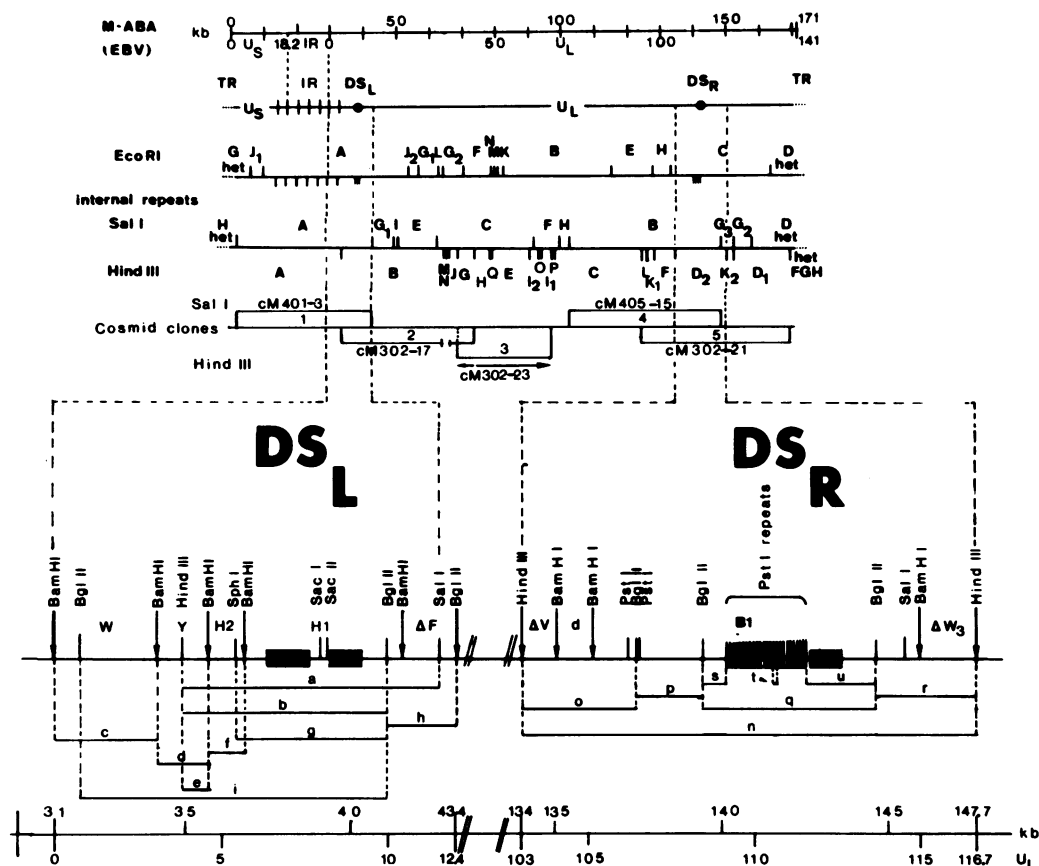


FIG. 1. Linkage of *EcoRI*, *SalI*, and *HindIII* fragments of the M-ABA (EBV) strain and an expanded fine map of the DS_L and DS_R regions. The map positions of the cosmid clones used as ^{32}P -labeled probes are indicated with open bars. The map positions of all other clones used for hybridization are indicated as solid lines below the expanded fine map of DS_L and DS_R . The small letters marking the clones were used in Fig. 2 through 7 and Table 1 to designate the labeled probes. A comprehensive description of the cloned probes used in these experiments is in preparation (A. Polack, G. Hartl, U. K. Freese, U. Zimmer, L. Gissmann, and G. W. Bornkamm, manuscript in preparation).

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, 70).

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. Zimmer, 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 μ g 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

$\times 10^6$ to 0.2×10^6 cells per ml of culture and harvested at 0.3×10^6 to 0.4×10^6 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 $0.1 \mu\text{m}$ (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 [³²P]dCTP (400 Ci/mmol; Amersham Corp., Arlington Heights, Ill.) in a total volume of 20 to 50 μl (45, 57). The specific activity of labeled probes was between 5×10^7 and 10^8 cpm/ μg 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. rRNAs (28S, 5,100 bases [b]; 18S, 1,800 b) served as internal size standards (11). Additionally, a DNA marker mix was used containing linear pHC79 dimer and monomer, pBR322 linear monomer, and additional digests of pBR322 with *Bam*HI-*Pst*II and with *Hin*II. 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 *Not*I (R. Borsetti, D. Wise, and T. Atkinson, unpublished data; 58) recognizes the 10-b palindrome GGC[↓]GGCCGCC, that the internal hexanucleotide palindrome is an *Xma*III site, that not all *Xma*III sites are substrates for *Not*I, and that not all *Not*I sites are substrates for *Xma*III. 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 ³²P-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. Zimmer, 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 μg) from induced (Fig. 2) and uninduced Raji cells was run on a denaturing gel (63), transferred to nitrocellulose (68), and hybridized with ³²P-labeled M-ABA viral DNA 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-

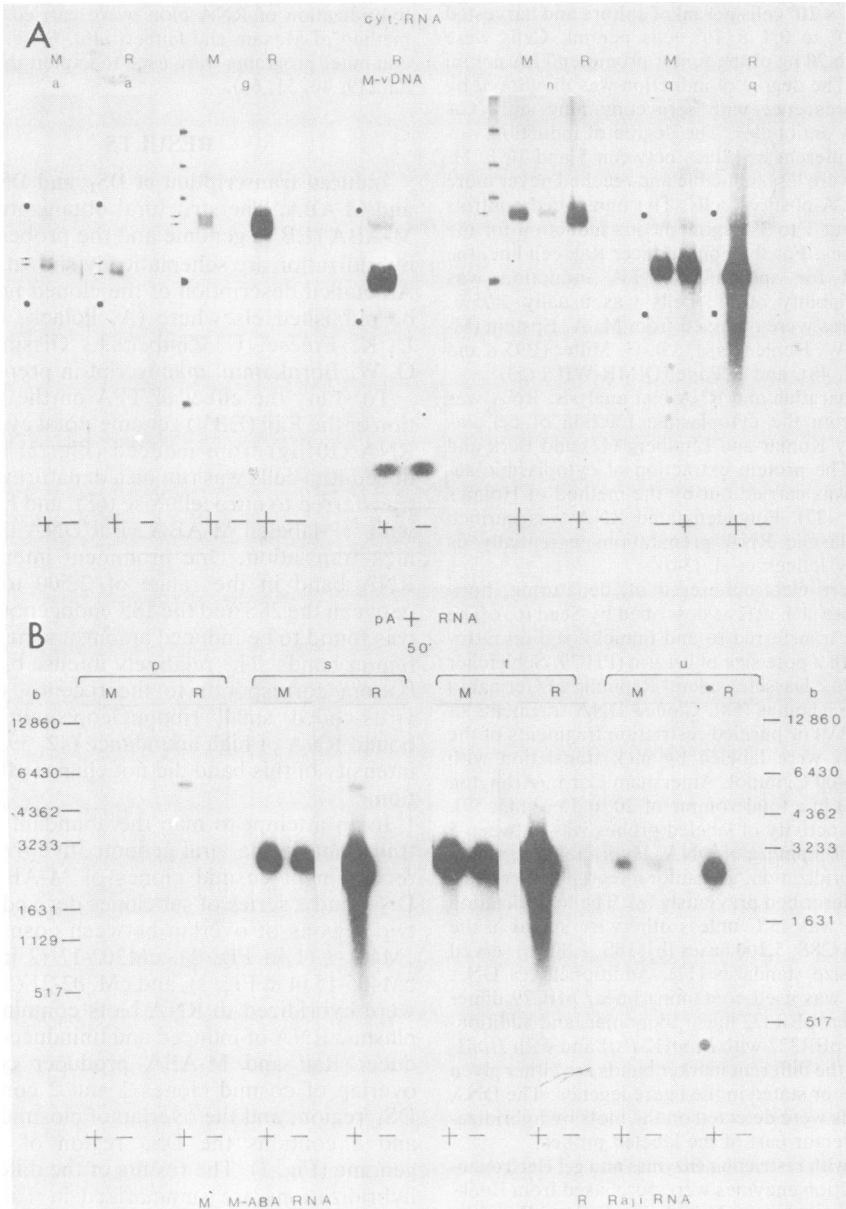


FIG. 2. (A) Autoradiograms of blots containing size-fractionated cytoplasmic RNAs (10 μ g per lane) of chemically induced (+) and uninduced (-) Raji (R) and M-ABA (M) cells after hybridization with 32 P-labeled M-ABA virus DNA (M-vDNA) and with different recombinant plasmid clones containing DS_L (probes 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 ($\approx 1 \mu$ g) which is contained in 10 μ g of unfractionated cytoplasmic RNA. Hybridization temperature was 50°C. Molecular weights (in b) are indicated at the sides of the outer blots.

ing region of DS_L are represented in the transcript. Correspondingly, within clone q (Fig. 2A) the sequences flanking the DS_R region are too short to code for an RNA of the respective size.

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

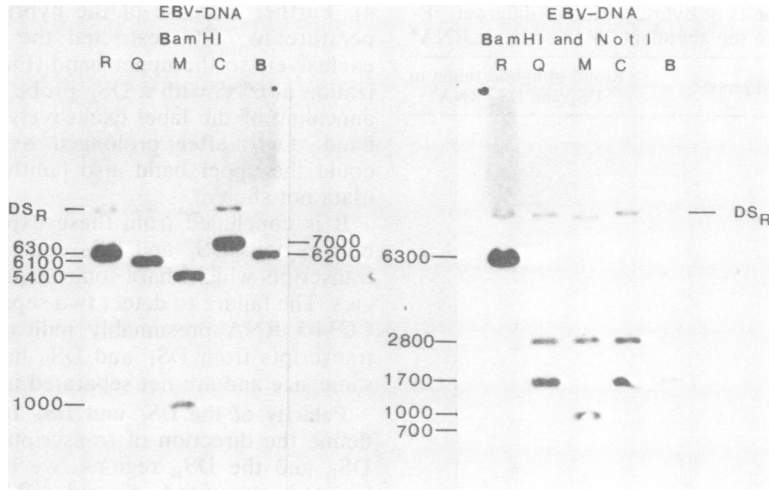


FIG. 3. Cellular Raji (R) and viral DNAs of the strains QIMR-WIL (Q), M-ABA (M), CC34-5 (C), and B95-8 (B) were digested with either *Bam*HI alone or double digested with *Bam*HI and *Not*I. The DNAs were fractionated on a 0.5% agarose gel and transferred to nitrocellulose. A ³²P-labeled DS_L-specific probe (probe g in Fig. 1) was used for hybridization. The size polymorphism of *Bam*HI fragment H of QIMR-WIL, M-ABA, CC34-5, and B95-8 (EBV) is lost upon double digestion with *Bam*HI and *Not*I, indicating that different numbers of *Not*I repeats are responsible for the size polymorphism in *Bam*HI-H. M-ABA (EBV) has an additional *Bam*HI cleavage site, giving rise to a fragment neighboring the repeats of smaller size (about 700 bp). The *Not*I repeats were not transferred to the nitrocellulose. Note that *Not*I does not cut within the *Bam*HI H fragment of Raji (EBV) DNA. All fragment sizes are given in bp.

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 *Pst*I site of pHCT9 (probe t in Fig. 1). The two *Bgl*III-*Pst*I 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 *Bgl*III-*Pst*I 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

TABLE 1. Results of hybridization of different ³²P-labeled probes to the abundant TPA-inducible RNA^a

Code of ³² P-labeled probe	Result of hybridization to DS _L and DS _R RNA ^b
Cosmids	
1	+
2	+
3	-
4	+
5	+
Left overlap	
a	+
b	+
c	-
d	-
e	-
f	-
g (DS _L)	+
h	-
Right overlap	
n	+
o	-
p	-
q (DS _R)	+
r	-
s ^c	+
t	+
u ^c	+

^a Map orientation of the probes used for hybridization are described in Fig. 1.

^b +, Hybridization positive; -, lack of hybridization.

^c Not cloned. Probe was prepared as a DNA fragment from an agarose gel.

DS_L, 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_R 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_L 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_L 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.

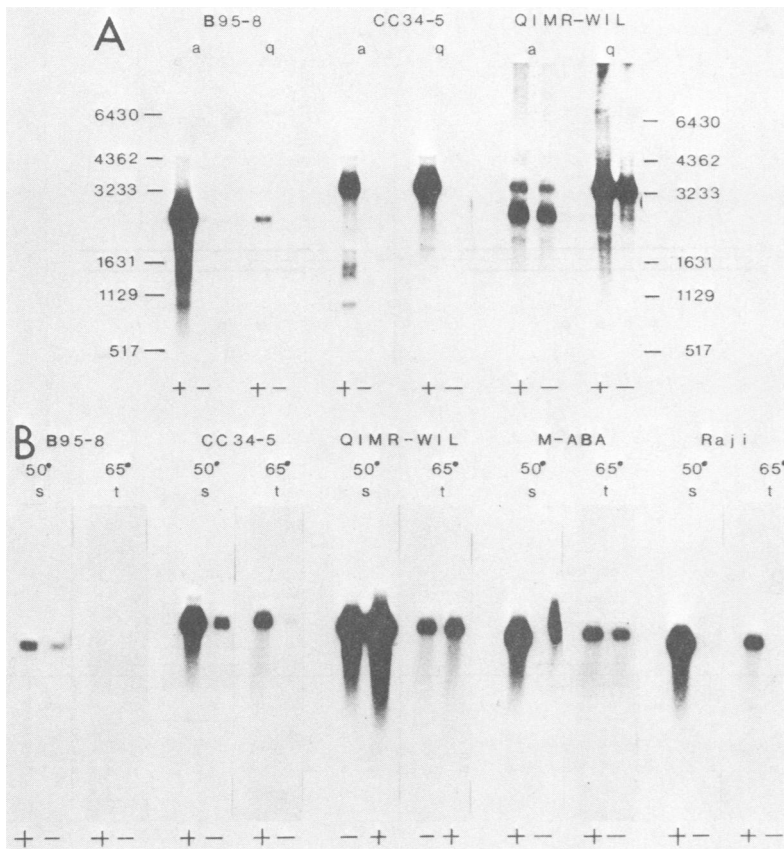


FIG. 4. Hybridization of size-fractionated RNA from five EBV strains with probes from DS_L and DS_R at conditions of different stringency. (A) Autoradiograms of size-fractionated total cytoplasmic RNAs from the additional virus strains B95-8, CC34-5, and QIMR-WIL hybridized at 45°C with the indicated probes. (B) Assignment of individual transcripts of five different strains to DS_R and DS_L by hybridization at 50°C (moderate stringency) and 65°C (high stringency). All ³²P-labeled hybridization probes are described in Fig. 1.

It has been shown that fragments carrying DS_L and DS_R exhibit a marked length polymorphism (7, 8, 56) and that different numbers of DS_R repeats account for the length polymorphism in fragments containing DS_R (29). Viral DNAs were digested either with *Bam*HI alone or with *Bam*HI and *Not*I, which cuts within the DS_L repeats (Fig. 5). The sizes (in b) of the *Bam*HI H fragments carrying DS_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 *Not*I 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 *Bam*HI cleavage site within *Bam*HI-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 *Bam*HI H fragments were pronounced after

*Bam*HI digestion alone. This indicates that different numbers of DS_L repeats account for the size variation in *Bam*HI-H. In contrast, the DS_L repeats of Raji (EBV) appeared not to be cut by *Not*I. The recognition sequence of *Not*I contains two CpG dinucleotides which are the main targets for methylation in higher cells. The fact that *Not*I does not cut within the DS_L repeats of Raji (EBV) might be due to extensive methylation of Raji (EBV) DNA (38).

The lengths (in b) of the corresponding DS_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 DS_L and DS_R repeats in each strain.

DNA sequence of the DS_R repeat. To study whether the DS_R repeats can possibly code for a protein, we determined the nucleotide sequence

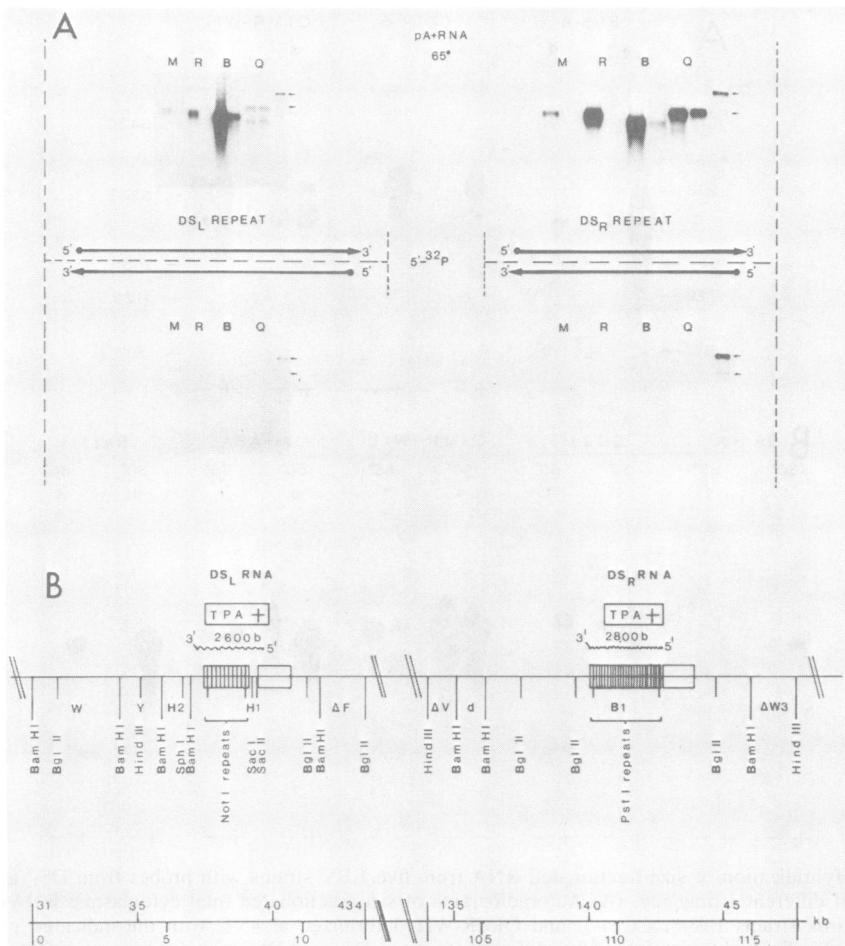


FIG. 5. (A) Autoradiograms of a blot hybridization experiment with separated ^{32}P -labeled single strands of the DS_L *NotI* repeat (DS_L REPEAT) and the DS_R *AvaI* repeat (DS_R REPEAT) fragments. The blots contain polyadenylated RNAs ($\approx 5 \mu\text{g}$ per lane) from induced (left of each pair of lanes) and uninduced (right of each pair of lanes) M-ABA (M), Raji (R), B95-8 (B), and QIMR-WIL (Q) cells. Hybridization was carried out at 65°C . The marker lanes at the far right of each panel were separately hybridized by a nick-translated cloned probe containing pHCT9 sequences at 42°C . Arrows at the sides of the autoradiograms indicate 4.36- and 3.23-kb marker positions. In the diagram, strand separation is symbolized by the horizontal dashed lines between the two pairs of DS_L and DS_R single strands. The abundance of the chemically induced DS_L RNA in B95-8 cells in this experiment is higher as in the other cell lines. For this reason, the differential hybridization of the DS_L and DS_R single strands can be better observed in the two lanes containing spontaneously induced B95-8 RNA. (B) Summary diagram of gene organization and transcription of the DS_L and DS_R regions. Only the length of the M-ABA DS_L (2,600 b) and DS_R (2,800 b) transcripts are included in the figure. The lengths of those RNAs from the other EBV strains are given in the text. This model is mainly based on the map data (Fig. 2 and Table 1), the differential hybridization data at different stringencies (Fig. 4 and Fig. 5A), the hybridization data of the strand-specific repeat probes (Fig. 5A), and the corresponding sizes of DS_L and DS_R RNAs with the lengths of the respective repeat clusters (Fig. 3 and 4) in all strains tested.

of one DS_R 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 DS_R 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,

this result was confirmed by sequencing the upper 3'-end-labeled strand from the nearby *Pst*I 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 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 polyadeny-

lated 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'  CGGCTGC AGC CGGCCACCCCCGCC GGAGCGGG CAGCGGACCCAGCGGACCC GGTGGCCACCC GGC 3'
#2  5'  CGG TGCCCAGCGGGGCCACCCCCACCC GGAGCGGG CAGCGG CCCAGCGGACCCACCCG CGCC GCCCCGC 3'

```

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

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

```

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

```

Prevalent sequences of S_{γ1}, S_{γ2b} and S_{γ3} regions (#4)

```

#4  Y1  5'  GP PT CC AGGCTG AGCAGCTA CA GG GGAGC T GGG G YA PPTGGGAPTTP G 3'
#4  Y2b 5'  GG GA CC AG7CCT AGCAGCTP TG GG GGAGC T GGG G A7 GGTGGGAPTTPG A 3'
#4  Y3  5'  PG NA CC 7GPNTP AGCAPYYA CA GG GGAGC T GGG G 7P GGTGGGAGTATP P 3'

```

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

```

#5  5'  GGAGCTGGG TGAGCTGGG TGAGCTGGG GGAGCTGGG GGAGCTGGG TGAGCTGGG GGAGCTGGG 3'

```

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 eucaryotic 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 motif, GGAGCGGG, of these 76 bp was also detected in the I_R-T_R inversion sequence of herpes simplex virus (16, 49) (Fig. 7). A short related sequence, GGAGC₆GGG, was also found in the switch site consensus sequences of the Ig-γ-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. Kofler 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

1. Arrand, J. R., and L. Rymo. 1982. Characterization of the major Epstein-Barr virus-specific RNA in Burkitt lymphoma-derived cells. *J. Virol.* 41:376-389.
2. Arrand, J. R., L. Rymo, J. E. Walsh, E. Björck, T. Lindahl, and B. E. Griffin. 1981. Molecular cloning of the complete Epstein-Barr virus genome as a set of overlapping restriction endonuclease fragments. *Nucleic Acids Res.* 9:2999-3014.
3. Artavanis-Tsakonas, S., P. Schedl, M. Mirault, L. Moran, and J. Lis. 1979. Genes for the 70,000 Dalton heat shock protein in two cloned *D. melanogaster* DNA segments. *Cell* 17:9-18.
4. Barker, W. C., L. K. Ketcham, and M. O. Dayhoff. 1978. A comprehensive examination of protein sequences for evidence of internal gene duplication. *J. Mol. Evol.* 10:265-281.
5. Bauer, G., P. Höfler, and H. zur Hausen. 1982. Epstein-Barr virus induction by a serum factor. I. Induction and cooperation with additional inducers. *Virology* 121:184-194.
6. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
7. Bornkamm, G. W., H. Delius, U. Zimmer, J. Hudewentz, and M. A. Epstein. 1980. Comparison of Epstein-Barr virus strains of different origin by analysis of the viral DNAs. *J. Virol.* 35:603-618.
8. Bornkamm, G. W., J. Hudewentz, U. K. Freese, and U. Zimmer. 1982. Deletion of the nontransforming Epstein-Barr virus strain P3HR-1 causes fusion of the large internal repeat to the DS_L region. *J. Virol.* 43:952-968.
9. Brutlag, D. L., J. Clayton, P. Friedland, and L. H. Keddes. 1982. SEQ: a nucleotide sequence analysis and recombination system. *Nucleic Acids Res.* 10:279-294.
10. Buell, G. N., D. Reisman, C. Kintner, G. Crouse, and B. Sugden. 1981. Cloning overlapping DNA fragments from the B95-8 strain of Epstein-Barr virus reveals a site of homology to the internal repetition. *J. Virol.* 40:977-982.
11. Chandler, P. M., D. Rimkus, and N. Davidson. 1979. Gel electrophoretic fractionation of RNAs by partial denaturation with methylmercuric hydroxide. *Anal. Biochem.* 99:200-206.
12. Crawford, D. H., M. A. Epstein, G. W. Bornkamm, B. G. Achong, S. Finerty, and J. L. Thompson. 1979. Biological and biochemical observations on isolates of EB virus from the malignant epithelial cells of two nasopharyngeal carcinomas. *Int. J. Cancer* 24:294-302.

13. Dambaugh, T., C. Beisel, M. Hummel, W. King, S. Fennewald, A. Cheung, M. Heller, N. Raab-Traub, and E. Kieff. 1980. Epstein-Barr virus (B95-8) DNA. VII. Molecular cloning and detailed mapping. *Proc. Natl. Acad. Sci. U.S.A.* 77:2999-3003.
14. Dambaugh, T. R., and E. Kieff. 1982. Identification and nucleoside sequences of two similar tandem direct repeats in Epstein-Barr virus DNA. *J. Virol.* 44:823-833.
15. Dambaugh, T., F. K. Nkrumah, R. J. Biggar, and E. Kieff. 1979. Epstein-Barr virus RNA in Burkitt tumor tissue. *Cell* 16:313-322.
16. Davison, A. J., and N. M. Wilkie. 1981. Nucleotide sequences of the joint between the L and S segments of herpes simplex virus types 1 and 2. *J. Gen. Virol.* 55:315-331.
17. Epstein, M. A., and B. G. Achong (ed.). 1979. The Epstein-Barr virus, p. 1-22. Springer Verlag, Berlin.
18. Epstein, M. A., and B. G. Achong (ed.). 1979. The Epstein-Barr virus, p. 61-78. Springer Verlag, Berlin.
19. Fisher, D. K., G. Miller, L. Gradoville, L. Heston, M. W. Weststrate, W. Maris, J. Wright, J. Brandsma, and W. C. Summers. 1981. Genome of a mononucleosis Epstein-Barr virus contains DNA fragments previously regarded to be unique to Burkitt's lymphoma isolates. *Cell* 24:543-553.
20. Geraghty, D., M. A. Peifer, I. Rubenstein, and J. Messing. 1981. The primary structure of a plant storage protein: zein. *Nucleic Acids Res.* 9:5163-5174.
21. Given, D., and E. Kieff. 1979. DNA of Epstein-Barr virus VI. Mapping of the internal tandem reiteration. *J. Virol.* 31:315-324.
22. Hayward, S. D., and E. D. Kieff. 1976. Epstein-Barr virus-specific RNA I. Analysis of viral RNA in cellular extracts and in the polyribosomal fraction of permissive and nonpermissive lymphoblastoid cell lines. *J. Virol.* 18:518-525.
23. Hayward, S. D., S. G. Lazarowitz, and G. S. Hayward. 1982. Organization of the Epstein-Barr virus DNA molecule II. Fine mapping of the boundaries of the internal repeat cluster of B95-8 and identification of additional small tandem repeats adjacent to the HR-1 deletion. *J. Virol.* 43:201-212.
24. Hayward, S. D., L. Noguee, and G. S. Hayward. 1980. Organization of repeated regions within the Epstein-Barr virus DNA molecule. *J. Virol.* 33:507-521.
25. Heller, M., A. Henderson, and E. Kieff. 1982. Repeat array in Epstein-Barr virus DNA is related to cell DNA sequences interspersed on human chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 79:5916-5920.
26. Heller, M., V. van Santen, and E. Kieff. 1982. Simple repeat sequence in Epstein-Barr virus DNA is transcribed in latent and productive infections. *J. Virol.* 44:311-320.
27. Holmes, D. S., and J. Bonner. 1973. Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. *Biochemistry* 12:2330-2338.
28. Hudewentz, J., G. W. Bornkamm, and H. zur Hausen. 1980. Effect of the diterpene ester TPA on Epstein-Barr virus antigen- and DNA synthesis in producer and nonproducer cell lines. *Virology* 100:175-178.
29. Hudewentz, J., H. Delius, U. K. Freese, U. Zimmer, and G. W. Bornkamm. 1982. Two distant regions of the Epstein-Barr virus genome with sequence homologies have the same orientation and involve small tandem repeats. *EMBO J.* 1:21-26.
30. Hummel, M., and E. Kieff. 1982. Epstein-Barr virus RNA VIII. Viral RNA in permissively infected B95-8 cells. *J. Virol.* 43:262-272.
31. Hummel, M., and E. Kieff. 1982. Mapping of polypeptides encoded by the Epstein-Barr virus genome in productive infection. *Proc. Natl. Acad. Sci. U.S.A.* 79:5698-5702.
32. Jat, P., and J. R. Arrand. 1982. In vitro transcription of two Epstein-Barr virus specified small RNA molecules. *Nucleic Acids Res.* 10:3407-3425.
33. Jeang, K.-T., and S. D. Hayward. 1983. Organization of the Epstein-Barr virus DNA molecule. III. Location of the P3HR-1 deletion junction and characterization of the *NotI* repeat units that form part of the template for an abundant 12-O-tetradecanoylphorbol-13-acetate-induced mRNA transcript. *J. Virol.* 48:135-148.
34. Jelinek, W., M. Adesnik, M. Salditt, D. Sheiness, R. Wall, G. Molloy, L. Philipson, and J. E. Darnell. 1973. Further evidence on the nuclear origin and transfer to the cytoplasm of polyadenylic acid sequences in mammalian cell RNA. *J. Mol. Biol.* 75:515-532.
35. Kataoka, T., T. Miyata, and T. Honjo. 1981. Repetitive sequences in class-switch recombination regions of immunoglobulin heavy chain genes. *Cell* 23:357-368.
36. King, W., V. van Santen, and E. Kieff. 1981. Epstein-Barr virus RNA VI. Viral RNA in restringently and abortively infected Raji cells. *J. Virol.* 38:649-660.
37. Kintner, C. R., and B. Sugden. 1979. The structure of the termini of the DNA of the Epstein-Barr virus. *Cell* 17:661-671.
38. Kintner, C., and B. Sugden. 1981. Conservation and progressive methylation of Epstein-Barr viral DNA sequences in transformed cells. *J. Virol.* 38:305-316.
39. Kramer, J. M., G. N. Cox, and D. Hirsh. 1982. Comparisons of the complete sequences of two collagen genes from *Caenorhabditis elegans*. *Cell* 30:599-606.
40. Kröger, M., and A. Kröger-Block. 1982. A flexible new computer program for handling DNA sequence data. *Nucleic Acids Res.* 10:229-236.
41. Kumar, A., and U. Lindberg. 1972. Characterization of messenger ribonucleoprotein and messenger RNA from KB cells. *Proc. Natl. Acad. Sci. U.S.A.* 69:681-685.
42. Lerner, M. R., N. C. Andrews, G. Miller, and J. A. Steitz. 1981. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. U.S.A.* 78:805-809.
43. Lindahl, T., A. Adams, G. Bjursell, G. W. Bornkamm, C. Kaschka-Dierich, and U. Jehn. 1976. Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line. *J. Mol. Biol.* 102:511-530.
44. Lizardi, P. M. 1979. Genetic polymorphism of silk fibroin studied by two-dimensional translation pause fingerprints. *Cell* 18:581-589.
45. Maniatis, T., A. Jeffrey, and D. G. Kleid. 1975. Nucleotide sequence of the rightward operator of phage lambda. *Proc. Natl. Acad. Sci. U.S.A.* 72:1184-1188.
46. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
47. Miller, G., D. Cooper, J. Niederman, and J. Pagano. 1976. Biological properties and viral surface antigens of Burkitt lymphoma- and mononucleosis-derived strains of Epstein-Barr virus released from transformed marmoset cells. *J. Virol.* 18:1071-1080.
48. Miller, G., and M. Lipman. 1973. Release in infectious Epstein-Barr virus by transformed marmoset leukocytes. *Proc. Natl. Acad. Sci. U.S.A.* 70:190-194.
49. Mocarski, E. S., and B. Roizman. 1981. Site-specific inversion sequence of the herpes simplex virus genome: domain and structural features. *Proc. Natl. Acad. Sci. U.S.A.* 78:7047-7051.
50. Moran, L., M. Mirault, A. Tissières, L. Lis, P. Schedl, S. Artavanis-Tsakonas, and W. J. Gehring. 1979. Physical map of two *D. melanogaster* DNA segments containing sequences coding for the 70,000 dalton heat shock protein. *Cell* 17:1-8.
51. Orcutt, B. C., D. G. George, J. A. Frederickson, and M. A. Dayhoff. 1982. Nucleic acid sequence database computer system. *Nucleic Acids Res.* 10:157-174.
52. Orellana, T., and E. Kieff. 1977. Epstein-Barr virus-specific RNA. II. Analysis of polyadenylated viral RNA in restringent, abortive, and productive infections. *J. Virol.* 22:321-330.
53. Pope, J. H., M. Horne, and W. Scott. 1968. Transformation of foetal leukocytes in vitro by filtrates of a human leukaemic cell line containing herpes-like virus. *Int. J. Cancer* 3:857-866.

54. **Pritchett, R. F., S. D. Hayward, and E. D. Kieff.** 1975. DNA of Epstein-Barr virus. I. Comparative studies of the DNA of Epstein-Barr virus from HR-1 and B95-8 cells: size, structure, and relatedness. *J. Virol.* **15**:556-569.
55. **Pulvertaft, R. J. V.** 1965. A study of malignant tumors in Nigeria by short term tissue culture. *J. Clin. Pathol.* **18**:261-273.
56. **Raab-Traub, N., T. Dambaugh, and E. Kieff.** 1980. DNA of Epstein-Barr virus. VIII. B95-8, the previous prototype, is an unusual deletion derivative. *Cell* **22**:257-267.
57. **Rigby, P. W., 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.
58. **Roberts, R. J.** 1982. Restriction and modification enzymes and their recognition sequences. *Nucleic Acids Res.* **10**:r117-r144.
59. **Rosa, M. D., E. Gottlieb, M. R. Lerner, and J. A. Steitz.** 1981. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. *Mol. Cell. Biol.* **1**:785-796.
60. **Rymo, L.** 1979. Identification of transcribed regions of Epstein-Barr virus DNA in Burkitt lymphoma-derived cells. *J. Virol.* **32**:8-18.
61. **Rymo, L., and S. Forsblom.** 1978. Cleavage of Epstein-Barr virus DNA by restriction endonucleases Eco RI, Hind III and Bam HI. *Nucleic Acids Res.* **5**:1387-1402.
62. **Sakoyama, Y., Y. Yaoita, and T. Honjo.** 1982. Immunoglobulin switch region-like sequences in *Drosophila melanogaster*. *Nucleic Acids Res.* **10**:4203-4214.
63. **Seed, B.** 1982. Diazotizable arylamine cellulose papers for the coupling and hybridization of nucleic acids. *Nucleic Acids Res.* **10**:1799-1810.
64. **Skare, J., and J. L. Strominger.** 1980. Cloning and mapping of BamHI endonuclease fragments of DNA from the transforming B95-8 strain of Epstein-Barr virus. *Proc. Natl. Acad. Sci. U.S.A.* **77**:3860-3864.
65. **Sprague, K. V., M. B. Roth, R. F. Manning, and P. L. Gage.** 1979. Alleles of the fibroin gene coding for proteins of different length. *Cell* **17**:407-413.
66. **Staden, R.** 1977. Sequence data handling by computer. *Nucleic Acids Res.* **4**:4037-4051.
67. **Sümeği, J., L. Wieslander, and B. Daneholt.** 1982. A hierarchic arrangement of the repetitive sequences in the Balbiani Ring 2 gene of *Chironomus tentans*. *Cell* **30**:579-587.
68. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* **77**:5201-5205.
69. **Thomas Powell, A. L., W. King, and E. Kieff.** 1979. Epstein-Barr virus-specific RNA. III. Mapping of DNA encoding viral RNA in restringent infection. *J. Virol.* **29**:261-274.
70. **van Santen, V., A. Cheung, and E. Kieff.** 1981. Epstein-Barr virus RNA. VII. Size and direction of transcription of virus-specified cytoplasmic RNAs in a transformed cell line. *Proc. Natl. Acad. Sci. U.S.A.* **78**:1930-1934.
71. **Yamada, Y., V. E. Avvedimento, M. Mudryj, H. Ohkubo, G. Vogeli, M. Irani, I. Pastan, and B. de Crombrughe.** 1980. The collagen gene: evidence for its evolutionary assembly by amplification of a DNA segment containing an Exon of 54 bp. *Cell* **22**:887-892.
72. **Yčas, M.** 1976. Origin of periodic proteins. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**:2139-2140.
73. **zur Hausen, H., G. W. Bornkamm, R. Schmidt, and E. Hecker.** 1979. Tumor initiators and promoters in the induction of Epstein-Barr virus. *Proc. Natl. Acad. Sci. U.S.A.* **76**:782-785.
74. **zur Hausen, H., F. J. O'Neill, U. K. Freese, and E. Hecker.** 1978. Persisting oncogenic herpesvirus induced by tumor-promoter TPA. *Nature (London)* **272**:373-375.