Haematology and Blood Transfusion Vol. 28
Modern Trends in Human Leukemia V
Edited by Neth, Gallo, Greaves, Moore, Winkler

Structural Organization and Expression of the Epstein-Barr Virus Genome*

G. W. Bornkamm, U. K. Freese, G. Laux, J. Hudewentz, and H. Delius

Infection of B-lymphocytes with Epstein-Barr virus (EBV) induces unlimited growth of the cells in culture. Cells immortalized by EBV harbor the viral genome and express the virus-specific nuclear antigen EBNA.

In EBV-carrying cell lines usually a small and varying percentage of cells is spontaneously induced to produce viral particle (producer lines). Treatment of the cells with the tumor promoter 12-0-tetradecanoyl-phorbol-13-acetate (TPA) and a number of other inducers increases the number of virus-producing cells. In some cell lines TPA induces only the synthesis of early viral antigens (nonproducer lines) [6].

The EBV genome isolated from purified virus particles is a linear double-stranded DNA molecule of 180,000 bp, carrying identical repeats of about 400 bp at both termini. It is composed of a short unique region (U_S) of 15 kbp and a long unique region (U_L) of 130 kbp, separated from each other by an array of 3.1-kpb repeats in tandem orientation. The number of repeats varies between 6 and 12 among different virus isolates and is variable even within one DNA population.

Sequences located in *Bam HI-H* at the left hand side of the long unique region have been shown to share some sequence homology with sequences about 105 kbp apart in *Bam HI-B1* at the right hand side of the long unique region [4]. The regions with sequence homologies are denoted DS_L and DS_R (left and right duplicated sequence). DS_R is located within the 12 kbp, which are delated in the B 95-8 strain of EBV.

The observation that DS_L and DS_R are actively transcribed after treatment of the cells with TPA led us to a more detailed analysis of the two regions with sequence homologies and their transcription products.

Recombinant plasmids containing DS_L and DS_R were obtained by cloning the DNA of M-ABA (EBV), a virus, originally derived from a nasopharyngeal carcinoma, which was shown to have the genomic organization of the EBV prototype [1]. Cleavage of the DNA of a clone containing DS_R with a number of different restriction endonucleases revealed a so far unrecognized small tandem repeat of 102 bp. Heteroduplexes of the DNA of the two clones containing DS_L and DS_R, respectively, visualized in the electron microscope by cytochrome c spreading, revealed that the region of homology is about 2.5 kbp long, involves the small repeats, and has the same orientation in the viral genome. Mica adsorption of the heteroduplex without cytochrome c showed that the region with sequence homology consists of about 1.5 kbp with partial homology including the small repeats and 0.9 kbp with well-matched duplexes. The involvement of the DS_R repeats in the heteroduplex formation suggested the presence of repeats also in DS_L. Reannealing of DNA containing the DS_L region can give rise to the formation of two loops of the same size, which is expected for a DNA containing tandem sequence reiterations. The presence of repeats in DS_L,

* This work was supported by the Deutsche Forschungsgemeinschaft (SFB 31)

including a restriction site for *NotI*, has recently been reported [2].

In the nonproducer cell line Raji transcription of DS_L and DS_R is switched on after TPA treatment. For the analysis of viral transcripts, cytoplasmic RNA was isolated from Raji and M-ABA cells treated with and without TPA and was run in formaldehyde agarose gels. The RNA was transferred from the gel to nitrocellulose according to Thomas [5]. Hybridization with ^{32}P-labeled cloned viral DNA fragments revealed that the RNA species coded for by DS_L and DS_R are 2.8 and 2.6 kb long. They are the most abundant transcripts after TPA induction. Hybridization with subclones showed that both RNA species are almost exclusively transcribed from the DS_L and DS_R repeats, respectively. Both RNA species are polyadenylated. Sequencing of the DS_R repeat revealed six open reading frames.

Whether the RNAs coded for by DS_L and DS_R are translated into polypeptides with a presumably repetitive amino acid sequence remains to be elucidated.

References

1. Bornkamm GW, Delius H, Zimber U, Hudewentz J, Epstein MA (1980) J Virol 35: 603–618
2. Hayward SD, Lazarowitz SG, Hayward GS (1982) J Virol 43:201–212
3. Hudewentz J, Delius H, Freese UK, Zimber U, Bornkamm GW (1982) The EMBO Journal 1:21–26
4. Raab-Traub N, Dambaugh T, Kieff E (1980) Cell 22:257–267
5. Thomas PS (1980) Proc Natl Acad Sci USA 73:5201–5205
6. Zur Hausen H, O'Neill FJ, Freese UK, Hecker G (1978) Nature 272:373–375