VIROLOGY 195, 578–586 (1993)

The Epstein–Barr Virus Candidate Vaccine Antigen gp340/220 Is Highly Conserved between Virus Types A and B[1]

JANICE F. LEES,[2] JANE E. ARRAND, STUART DE V. PEPPER, JAMES P. STEWART, MICHAEL MACKETT, AND JOHN R. ARRAND[3]

Cancer Research Campaign Laboratories, Paterson Institute for Cancer Research, Christie Hospital NHS Trust, Manchester M20 9BX, United Kingdom

Received January 5, 1993; accepted April 7, 1993

Anti-Epstein–Barr Virus (EBV) vaccines are being developed which are based on the gp340/220 membrane antigen (MA) gene products from the B95-8 strain. Some proteins are known to be immunologically quite different between type-A (1) and type-B (2) strains of EBV and therefore from a vaccine point of view it was critical to evaluate the degree *of conservation of gp340/220. The complete MA coding sequence was determined for two B-type viruses, AG876 and* P3HR-1, for comparison with the A-type B95-8. A variable region within MA was sequenced from several other strains. In addition the other open reading frames within the MA-containing *Bam*HI-L fragment of AG876 were sequenced and *compared. The results show that there is a high degree of homology between all strains examined. Although some* differences were found within the MA coding sequence the only major site of variation was within the repeat region and no consistent A/B changes were found. Monoclonal antibodies generated against A-type MA and representing six *epitope groups along the length of the gp340 molecule were found to recognize B-type gp340, thereby demonstrating* functional homology. We conclude that, as a vaccine antigen, B95-8 gp340/220 should be equally effective against both types of EBV. © 1993 Academic Press, Inc.

INTRODUCTION

Epstein–Barr virus (EBV), a member of the herpesvirus group, is the causative agent of infectious mononucleosis and may play an etiological role in the development of Burkitt's lymphoma and nasopharyngeal carcinoma (reviewed by de Thé, 1985). In addition, the virus is associated with the appearance of lymphomas in immunosuppressed patients (Thomas and Crawford, 1989) and seems to be involved in cases of Hodgkin's disease (Wu *et al.*, 1990). The introduction of a vaccine effective against EBV is therefore an important objective. To date vaccine development efforts have centered upon the use of the EBV major envelope glycoprotein gp340/220 (membrane antigen, MA) (reviewed in Arrand, 1992) which mediates virus adsorption and penetration via the complement receptor molecule CR2 [CD21] (Nemerow *et al.*, 1987; Tanner *et al.*, 1987) and is a major target for the production of virus-neutralizing antibodies (Thorley-Lawson and Poodry, 1982). As a result of these efforts, initial small-scale trials of a vaccinia recombinant-expressing membrane antigen have begun in China (Gu *et al.*, 1991) while development of a potential source of the glycoproteins for a subunit vaccine is well advanced (Madej *et al.*, 1992) and clinical trials are planned for the near future.

The efficacy of the response to a vaccine immunogen is clearly influenced by antigenic variation within the target protein. The overall extent of such genetic variation between different isolates of EBV is uncertain. Restriction enzyme analyses have revealed deletions, fragment polymorphisms, and differences in the extent of repeat sequences (Lung *et al.*, 1988; Bornkamm *et al.*, 1980, 1984; Dambaugh *et al.*, 1980; Heller *et al.*, 1981; Harris *et al.*, 1984; Rymo *et al.*, 1979, 1981; Sugden, 1977), although this level of variability is consistent with that expected for wild type isolates of any herpesvirus. However, two distinct EBV types (types A and B), sometimes referred to as EBV-1 and EBV-2, have been defined on the basis of specific sequence variation within the EBNA-2 gene encoding antigenically distinct forms of EBNA-2 (Adldinger *et al.*, 1985; Dambaugh *et al.*, 1984; Zimber *et al.*, 1986). More recent studies have also revealed that such type-specific differences extend to the EBNA-3 a,b,c genes and to the transcription units of the EBER RNAs (Rowe *et al.*, 1989; Sample *et al.*, 1990; Arrand *et al.*, 1989).

Since to date all potential vaccines have been based on gp340/220 from the B95-8 A-type strain, there existed the possibility that if the glycoproteins were antigenically distinct, a B-type virus may evade the effects

[1] The nucleotide sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under the Accession Numbers L07922 for strain P3HR-1 and L07923 for strain AG876.

[2] Current address: Department of Biochemistry and Molecular Biology Univ. of Manchester, Manchester M13 9PT, United Kingdom.

[3] To whom correspondence and reprint requests should be addressed.

0042-6822/93 $5.00

of such a vaccine. We decided to address the question of inter- and intrastrain variation within the MA gene first by nucleotide sequence analysis of two type-B viruses, P3HR-1 and AG876, and subsequent comparison with data available for the prototype B95-8 EBV strain (type-A), and second by examination of the ability of a panel of gp340/220-specific monoclonal antibodies to recognize cells containing A-type or B-type EBV.

MATERIALS AND METHODS

Plasmids and DNA sequencing

The MA gene is located within the *Bam*HI-L fragment of the EBV genome. This fragment from strain AG876, subcloned into pBluescript KS+ (Stratagene), was a gift from Dr. Jeffery Sample while the P3HR-1 putative *Bam*HI-L fragment was isolated from cosmid clone CP101.41 (kindly provided by Dr. Georg Bornkamm) and inserted into pBluescript pSK+. Sets of nested deletions were generated using a Pharmacia double-stranded nested deletion kit according to the manufacturer's instructions, with the exception of a Klenow end-fill reaction which was performed immediately prior to ligation and transformation. The nucleotide sequence was determined using the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) and modified T7 polymerase (United States Biochemical Corporation). Sequence analysis of deletion clones involved the use of small-scale plasmid preparations (Sambrook *et al.*, 1989) and the M13 universal primer. Additional sequence data were obtained using *Bam*HI-L specific oligonucleotide primers. Sequences were assembled using the SAP programs of Staden (1987) and analyzed using the GCG suite of programs (Devereux *et al.*, 1984).

B95-8 nucleotide coordinates are those used in the Genbank Database (release 73.0) which are derived from the sequence of Baer *et al.* (1984).

Cell lines and preparation of genomic DNA

B-lymphoid cell lines B95-8 and FF41 are derived from American cases of infectious mononucleosis; C2.BL16 is a lymphoblastoid cell line established by infection of adult seronegative blood with the B-type EBV strain BL16 (Rowe *et al.*, 1989); M-ABA is from an African NPC; P3HR-1, Chep, Wan, BL16, BL29, BL74, and AG876 are from African Burkitt's lymphomas; Igunga, Alvoch, and Ojango are from normal African individuals (Young *et al.*, 1987); and Akata is from a Japanese BL. Du, Zhu, and Zhang are biopsies taken from Chinese cases of NPC (Chen *et al.*, 1992). Cell lines were maintained as suspension cultures in RPMI 1640 medium (Gibco) supplemented with 10% fetal bovine serum (SeraLabs) at 37° in an atmosphere containing 6% CO_2. Total cellular DNA was prepared as described by Sambrook *et al.* (1989).

DNA amplification

DNA amplification was performed essentially as described by Lees *et al.* (1992). Oligonucleotide primers were prepared by the phosphoramidite method on an Applied Biosystems 392 synthesiser. Amplification reactions were analyzed by electrophoresis on 4% composite agarose gels consisting of 3% low-melting-point agarose and 1% normal agarose.

The nucleotide sequence of the portion of the gp340 sequence which was missing from cosmid CP101.41 (see Results) was obtained by PCR amplification of P3HR-1 total cellular DNA using the primers GGGACGTGAAGCAAAGAA (B95-8 coordinates 89,377–89,394) and ACTGCAGTGGGCCTCTCT (reverse complement of 89,548–89,565), in which one of the oligonucleotide primers was biotinylated, followed by DNA strand separation using streptavidin-coated magnetic beads (Dynal) and direct nucleotide sequencing as described above.

To investigate the repeat region of the gp340 coding sequence, total cellular DNA was amplified using the primers TCCTAATGTGTGGTTGGT (B95-8 coordinates 90,261–90,278) and GTCAGGCGCATCACCGGT (reverse complement of 90,690–90,708), which flank the repeat region, followed by strand separation and direct sequencing as above. In some cases the PCR product was cloned into plasmid pT7Blue (Novagen) prior to sequence analysis.

Immunofluorescence assays

Monoclonal antibodies against several epitopes of gp340 (Qualtiere *et al.*, 1987) were obtained from Dr. Louis Qualtiere, with the exception of antibody 72A1 (Hoffman *et al.*, 1980) which was provided by Dr. Diane Hayward. Monoclonal antibody OKT3, which recognizes the T-cell surface marker CD3, was obtained from ATCC. Live-cell immunofluorescence assays on B95-8 cells (type A) and C2.BL16 (type B) were performed according to Janossy and Amlot (1987). Cells were chemically induced to EBV late antigen production using 12-*O*-tetradecanoylphorbol-13-acetate (TPA) and 5-bromodeoxyuridine (BUdR) (Bauer, 1983). The proportions of MA-positive cells before and after induction were 1.4 and 29% for B95-8, and 0.8 and 17% for C2.BL16, respectively.

RESULTS

This work has determined the complete nucleotide sequence of the *Bam*HI-L fragment from EBV strain AG876 (4988 nucleotide pairs) and the sequence from the C-terminus of the gp340 gene to the right end of *Bam*HI-L from strain P3HR-1 (3210 nucleotide pairs).

The sequence analysis revealed that the putative *Bam*HI-L fragment subcloned from cosmid CP101.41 was a chimeric molecule, the center point of the rearrangement being a *Sal*I site. The structure of the resultant clone is such that the leftmost 1479 bp (*Bam*HI to *Sal*I) actually correspond to a different region of the EBV genome (coordinates 103,816 to 105,296 within the *Bam*HI-R fragment) while the remaining 3189 bp (*Sal*I to *Bam*HI) are derived from the *Bam*HI-L DNA fragment. This structure is a cloning artifact since the P3HR-1 library was constructed in a similar way to the M-ABA library (Polack *et al.*, 1984), from a partial *Sal*I digest of virus DNA. As a consequence of this rearrangement, the carboxy-terminal 52 nucleotides of the MA ORF were missing from the cosmid subclone. The nucleotide sequence of this region was obtained by PCR amplification of P3HR-1 total cellular DNA.

Examination of the gp340/220 ORF from the two B-type strains, AG876 and P3HR-1, predicted an identical polypeptide sequence from these two viruses which differed somewhat from that of the prototype B95-8 A-type. An alignment of the predicted MA amino acid sequences from B95-8 and P3HR-1/AG876 is shown in Fig. 1. The B95-8 MA ORF is 2721 bp in length encoding a protein of 907 aa while the equivalent ORFs in P3HR-1 and AG876 are only 2658 bp, predicting a protein of 886 aa. The size discrepancy is accounted for by the presence of two short deletions of 21 bp (7 aa) and 42 bp (14 aa) within the repeat region (see below). In addition to the deletions there are 37 base changes resulting in a total of 24 amino acid substitutions of which 6/24 were conservative changes (Fig. 1). The EBV MA protein is extensively glycosylated containing both N- and O-linked sugar moieties. The B95-8 sequence predicts a total of 36 potential N-linked glycosylation sites (Asn–X–Ser/Thr) while P3HR-1 and AG876 contain 35 such recognition sites. A summary of amino acid substitutions is provided in Fig. 5. Overall levels of identity are 98 and 97% at the nucleotide and amino acid levels, respectively.

In order to determine whether this high degree of sequence homology was conserved at a functional level, a panel of 14 different monoclonal antibodies raised against A-type EBV and representing six different epitope groups along the length of the gp340 molecule was used in immunofluorescence assays against homologous A-type B95-8 cells and heterologous B-type C2.BL16 cells. It was found that all six epitope groups were recognized in both cell lines indicating functional conservation of B-cell epitopes between gp340 from both types of EBV. Figure 2 shows a representative immunofluorescence and Table 1 summarizes the data for all monoclonal antibodies.

The MA polypeptide sequence contains a repetitive region which appears to consist of a basic 21-amino-acid repeat unit composed of three 7-amino-acid subunits. In B95-8 the repeat region is composed of a 14-amino-acid sequence repeated twice (aa 507–534), i.e., the first two repeat subunits, which expand into the full 21-residue segment repeated four times (aa 535–618). Molecular modeling studies suggest that this domain may interact with the phospholipid mem-

```
  1 MEAALLVCQYTIQSLIHLTGEDPGFFNVEIPEFPFYPTCNVCTADVNVTI  50
    ||||||||||||||||||:|| :|||||||| |||||.|||||||||.||
  1 MEAALLVCQYTIQSLIQLTRDDPGFFNVEILEFPFYPACNVCTADVNATI  50

 51 NFDVGGKKHQLDLDFGQLTPHTKAVYQPRGAPGGSENATNLFLLELLGAG 100
    |||||||||.|:|||| ||||||||||||||||||||||||||||||||||
 51 NFDVGGKKHKLNLDFGLLTPHTKAVYQPRGAPGGSENATNLPLLELLGAG 100

101 ELALTMRSKKLPINVTTGEEQQVSLESVDVYFQDVFGTMWCHHAEMQNPV 150
    ||||||||||||||:|||||||||||||||||||||||||||||||||||
101 ELALTMRSKKLPINITTGEEQQVSLESVDVYFQDVFGTMWCHHAEMQNPV 150

151 YLIPETVPYIKWDNCNSTNITAVVRAQGLDVTLPLSLPTSAQDSNFSVKT 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 YLIPETVPYIKWDNCNSTNITAVVRAQGLDVTLPLSLPTSAQDSNPSVKT 200

201 EMLGNEIDIECIMEDGEISQVLPGDNKFNITCSGYESHVPSGGILTSTSP 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 EMLGNEIDIECIMEDGEISQVLPGDNKFNITCSGYESHVPSGGILTSTSP 250

251 VATPIPGTGYAYSLRLTPRPVSRFLGNNSILYVFYSGNGPKASGGDYCIQ 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 VATPIPGTGYAYSLRLTPRPVSRFLGNNSILYVFYSGNGPKASGGDYCIQ 300

301 SNIVFSDEIPASQDMPTNTTDITYVGDNATYSVPMVTSEDANSPNVTVTA 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 SNIVFSDEIPASQDMPTNTTDITYVGDNATYSVPMVTSEDANSPNVTVTA 350

351 FWAWPNNTETDFKCKWTLTSGTPSGCENISGAPASNRTFDITVSGLGTAP 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 FWAWPNNTETDFKCKWTLTSGTPSGCENISGAPASNRTFDITVSGLGTAP 400

401 KTLIITRTATNATTTTHKVIFSKAPESTTTSPTLNTTGPADPNTTTGLPS 450
    ||||||||||||||||||||||||||||||||||||||.|||||||||||
401 KTLIITRTATNATTTTHKVIFSKAPESTTTSPTLNTTGPAAPNTTTGLPS 450

451 STHVPTNLTAPASTGPTVSTADVTSPTPAGTTSGASPVTPSPSPWDNGTE 500
    ||||||||||||||||||||||||||||||||||||||||||:|||||||
451 STHVPTNLTAPASTGPTVSTADVTSPTPAGTTSGASPVTPSPSPRDNGTE 500

501 SKAPDMTSSTSPVTTPTPNATSPTPAVTTPTPNATSPTPAVTTPTPNATS 550
    ||||||      ||||.|||||||||||||||||||||||||||||||||
501 SKAPDM..............TSPTSAVTTPTPNATSPTPAVTTPTPNATS 536

551 PTLGKTSPTSAVTTPTPNATSPTLGKTSPTSAVTTPTPNATSPTLGKTSP 600
    ||||||||||||||||||      ||||.|||||||| ||||||||||||
537 PTLGKTSPTSAVTTPTPNA.......TSPTPAVTTPTPNATIPTLGKTSP 579

601 TSAVTTPTPNATGPTVGETSPQANATNHTLGGTSPTPVVTSQPKNATSAV 650
    ||||||||||||:|||||||||||.|||||||||.||||||.|||||||||
580 TSAVTTPTPNATSPTVGETSPQANTTNHTLGGTSSTPVVTSPPKNATSAV 629

651 TTGQHNITSSSTSSMSLRPSSNPETLSPSTSDNSTSHMPLLTSAHPTGGE 700
    |||||||||||||||||||| .|||||||||||||||||||||||||||||
630 TTGQHNITSSSTSSMSLRPSSISETLSPSTSDNSTSHMPLLTSAHPTGGE 679

701 NITQVTPASISTHHVSTSSPAPRPGTTSQASGPGNSSTSTKPGEVNVTKG 750
    |||||||||.|||||||||||||||||||||||||||||||||||||||||
680 NITQVTPASTSTHHVSTSSPAPRPGTTSQASGPGNSSTSTKPGEVNVTKG 729

751 TPPQNATSPQAPSGQKTAVPTVTSTGGKANSTTGGKHTTGHGARTSTEPT 800
    |||.|||||||||||||||||||||||||||||||||||||||||||||||
730 TPPKNATSPQAPSGQKTAVPTVTSTGGKANSTTGGKHTTGHGARTSTEPT 779

801 TDYGGDSTTPRPRYNATTYLPPSTSSKLRPRWTFTSPPVTTAQATVPVPP 850
    ||||||||||.|||||||||||||||||||||||||||||||||||||||
780 TDYGGDSTTPRTRYNATTYLPPSTSSKLRPRWTFTSPPVTTAQATVPVPP 829

851 TSQPRFSNLSMLVLQWASLAVLTLLLLLVMADCAFRRNLSTSHTYTTPPY 900
    ||||||||||||||||||||||||||||||||||||||||||||||||||
830 TSQPRFSNLSMLVLQWASLAVLTLLLLLVMADCAFRRNLSTSHTYTTPPY 879

901 DDAETYV 907
    |||||||
880 DDAETYV 886
```

FIG. 1. An alignment of the predicted amino acid sequences of gp340 from EBV strains B95-8 (upper sequence) and P3HR-1/AG876 (lower sequence) generated using the BESTFIT program. The output shows amino acid homology on the basis of the evolutionary distance between amino acids. (|), Amino acid identity; (:), close relationship; (.), distant relationship; blank, no homology.

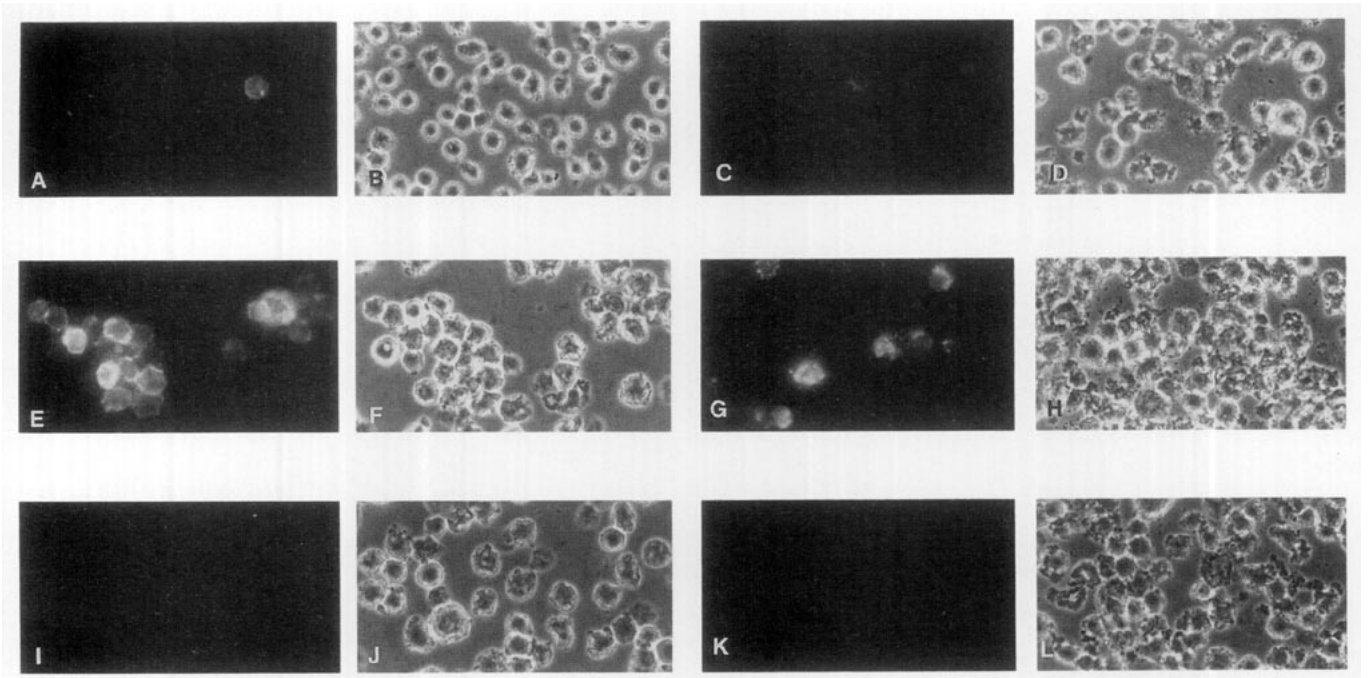


FIG. 2. gp340/220 immunofluorescence analysis of B95-8 and C2.BL16 cells. B95-8 cells (A, B, E, F, I, J) and C2.BL16 (C, D, G, H, K, L) either uninduced (A–D) or induced with TPA and BUdR (E–L) were assayed for expression of gp340/220 as described in the text. Fluorescence obtained using the gp340/220-specific monoclonal antibody F34 1F2 is shown in A, C, E, and G. The negative control antibody OKT3 was used in I and K. B, D, F, H, J, and L show phase contrast micrographs of the corresponding immunofluorescence fields. Magnification ×100.

brane (Beisel *et al.*, 1985). Analysis of the P3HR-1 and AG876 sequences indicates that, relative to the B95-8 prototype, the second 14-amino-acid segment has been deleted along with the last 7 amino acids of a 21-residue repeat (Figs. 1 and 4).

TABLE 1

REACTIVITY OF gp340/220-SPECIFIC MONOCLONAL ANTIBODIES WITH B95-8 (A-TYPE) AND C2.BL16 (B-TYPE) CELLS

Epitope group	Antibody	B95-8	C2.BL16
I	72A1*	+	+
	F30 3C2	+	+
	F34 4E3	+	+
	F29 1G7*	+	+
	F34 1F2*	+	+
	F34 5D3	+	+
	F34 6B1*	+	+
	F34 6B5	+	+
III	F16 3E3	+	+
IV	F29 89*	+	+
V	F16 C10	+	+
VI	F34 5H7	+	+
	F34 2B11	+	+
VII	F30 5C8	+	+
Negative control	OKT3	−	−

Note. Asterisk indicates a virus-neutralizing antibody. Note that different antibodies within the same epitope group do not necessarily recognize identical determinants as evidenced, for example, by different antibodies within epitope group 1 being either virus-neutralizing or nonneutralizing.

We have used PCR techniques to examine the structure of the MA repeat domain among a number of EBV type-A and type-B strains derived from different geographical regions and clinical sources. It was found that the length of the repeat region varied quite considerably between different strains as shown in Fig. 3. The PCR products of the repeat regions from a number of these strains were sequenced and found to differ in length from 399 base pairs (Alvoch) to 168 base pairs (Chep). The deduced amino acid sequences of six of these are shown in Fig. 4. They were all variants around the basic pattern of repeat units. In all strains the first repeat unit is only 14-amino-acids long, i.e., it lacks the third 7-residue subunit. Relative to the B95-8 sequence the differences consist of deletions (or in the case of Alvoch, insertion) of full 21-residue repeats or 7-residue third subunits (see Fig. 4). The significance of this is uncertain but it is consistent with strain variation noted within repetitive elements located elsewhere in the EBV genome (Lung *et al.*, 1988). It also provides a rational explanation for the differences in molecular weight estimates of MA that have been obtained in various laboratories.

The complete sequence of *Bam*HI-L from strain AG876 has been determined. In addition to the coding sequence for gp340/220, this fragment encodes the N-terminal exon of the latent protein EBNA 3a and several other polypeptides which are associated with productive infection. The extent of the homology of these sequences with their B95-8 counterparts is shown in

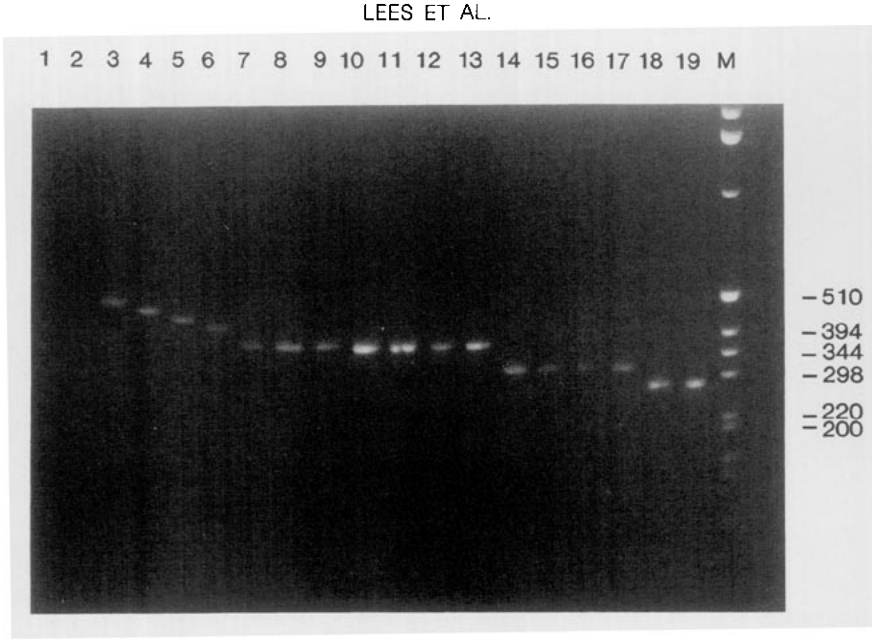


FIG. 3. PCR analysis of the repeat regions of the gp340 gene from different EBV strains. Total cellular DNA was amplified as detailed in the text. Products were fractionated by agarose gel electrophoresis and visualized under ultraviolet light. Lane 1, no DNA; lane 2, Ramos (EBV-negative); lane 3, Alvoch (type-B); lane 4, BL29 (type-B); lane 5, B95-8 (type-A); lane 6, BL16 (type-B); lane 7, P3HR-1 (type-B): lane 8, AG876 (type-B); lane 9, M-ABA (type-A); lane 10, FF41 (type-A); lane 11, Ojango (type-B); lane 12, Zhu (type-A); lane 13, BL74 (type-A); lane 14, Akata (type-A); lane 15, Du (type-A); lane 16, Igunga (type-A); lane 17, Zhang (type-A); lane 18, Chep (type-B); lane 19, Wan (type-B); Lane M contains size markers of the indicated sizes in base pairs.

Table 2. It is clear that in this region of the genome, interstrain sequence conservation is very high.

DISCUSSION

The variety in the extent of the repeat region predicts that in different strains, gp340 will exhibit substantial differences in size. This is due not only to the variation in length of the polypeptide but also because the high frequency of threonine and serine residues (potential O-linked glycosylation sites) in the repeat region and the existence of a potential N-linked glycosylation site within each 21-amino-acid repeat unit predicts a substantial effect on the total carbohydrate content of the glycoprotein. In contrast, gp220 is expected to be invariant between strains since the repeat region is spliced out of the mRNA for this molecule.

Figure 5 shows a schematic of gp340 indicating the domains of the molecule to which various relevant immunological functions have been mapped. It is notable that of the 24 amino acid changes between B95-8 and AG876/P3HR-1, 9 of these cluster in the region which is spliced out of gp220. The gp340 molecule contains a high proportion of serine, threonine, and proline residues, the majority of which are contained in the repeat region (Fig. 4). This unusual amino acid composition and variable extent of the repeat, together with the fact that it is removed from gp220, suggest that this portion of the molecule may not be essential for persistence and replication of fully functional virus.

Purified gp340/220 and recombinant vaccinia virus expressing gp340 have been employed successfully to protect cotton-top tamarins from a challenge dose of EBV (Finerty *et al.*, 1992; Morgan *et al.*, 1988). However, it has been noted that protection against EBV infection or virally induced disease does not always correlate with the induction of neutralizing antibodies (Epstein *et al.*, 1986; Emini *et al.*, 1989) suggesting that both antibody and cell-mediated responses contribute to protective immunity.

Sites of B-cell epitopes have been found along the length of the gp340 molecule (Pither *et al.*, 1992a,b) and three nonvirus-neutralizing monoclonal antibody binding sites have been mapped (Zhang *et al.*, 1991). Although at least two distinct virus-neutralizing epitopes are known to be present (Qualtiere *et al.*, 1987) their precise location is not yet determined. Monoclonal antibody 72A1 (Hoffman *et al.*, 1980) blocks EBV adsorption and belongs to an immunodominant murine monoclonal antibody group whose binding site is probably located within the NH_2-terminal 162 aa (Tanner *et al.*, 1988). Sairenji and co-workers (1988) demonstrated that 72A1 neutralizes both B95-8 and P3HR-1 viruses suggesting that the epitope is functional in both type-A and type-B viruses. The present study has confirmed the ability of antibody 72A1 to recognize both A-type and B-type gp340 and has also demonstrated conservation of several B-cell epitopes within these molecules.

Two CD4+ (helper) T-cell specific epitopes have been mapped by Wallace and co-workers (1991).

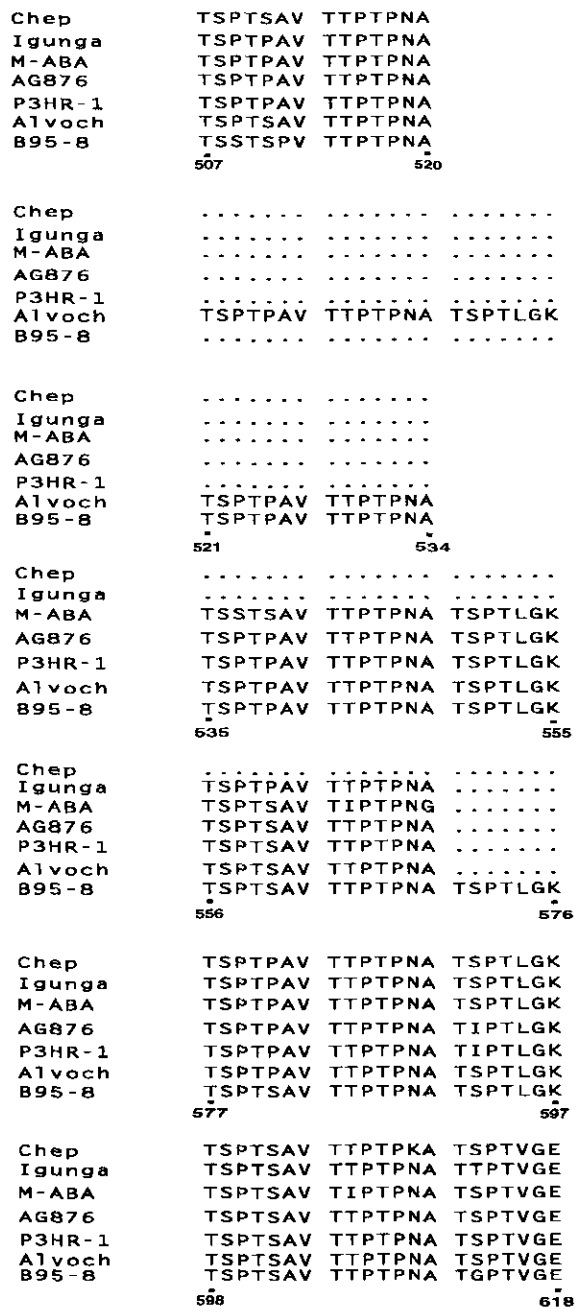


FIG. 4. Predicted amino acid sequences through the repeat region of gp340 from several strains of EBV. The numbers correspond to the B95-8 gp340 polypeptide sequence. (–) Indicates residues deleted relative to B95-8.

These are located between positions aa 61–81 and 163–183. The alignment between B95-8 and P3HR-1/AG876 sequences reveals a single nonconservative change Q(Gln) to L(Leu) at position 67. The CR2 binding site responsible for virus attachment and adsorption has been defined as EDPGFFNVE and spans aa 21–29. We note a single conservative substitution E(Glu) to D(Asp) within this domain.

The data obtained from this comparative study indicate that the EBV membrane antigen is highly conserved (>97% identity) between type-A and type-B strains. Many of the changes which do occur lie outside the currently defined domains of immunological significance. These results are in contrast to the type-specific variation detected within the latent protein genes EBNA2, EBNA3, a,b,c, where amino acid homology was reported to be 53, 84, 80, and 72%, respectively (Dambaugh *et al.*, 1984; Sample *et al.*, 1990). Few comparative sequence data are available comparing A-type and B-type viruses at other genetic loci although it has been reported that both the BZLF1 ORF and the IR1-derived coding exons for EBNA-LP of B95-8 and P3HR-1 are highly conserved (Jenson *et al.*, 1987; Jenson and Miller, 1988). In the course of this work we have obtained the complete sequence of AG876 *Bam*HI-L. This fragment contains several genes which exhibit a very high degree of conservation with the B95-8 homologues (Table 2).

Consequently, the pattern which is emerging shows two closely related wild-type virus strains which exhibit significant but specific divergence at several genetic loci. These variant loci appear to be confined to the latent genes, while proteins associated with productive infection are much more highly conserved. In some viral systems the major envelope glycoproteins are subject to genetic variation due to continued immunologic surveillance (Skehel and Wiley, 1986). The high level of sequence homology between EBV MA genes may be the result of structural/functional constraints. Alternatively, it may reflect the life-cycle of EBV in humans in which the virus latently infects B-lymphocytes and transcription from the majority of the viral genome (including gp350/220) is down-regulated (reviewed by Klein, 1989) thereby reducing immunologic selection pressure.

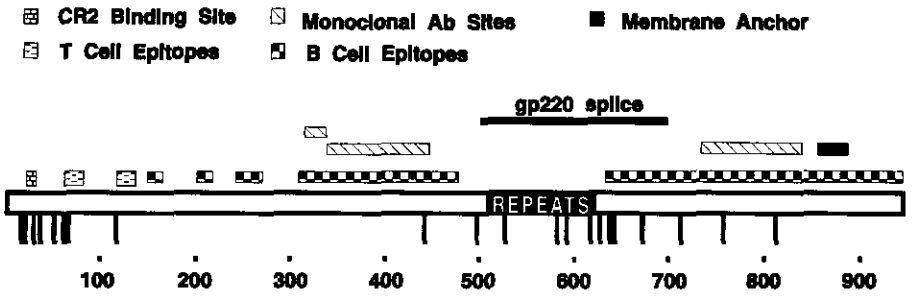
Although type-A EBV seems to be generally predominant in most parts of the world there do not appear to be disease-specific subtypes and it has been observed

TABLE 2

A COMPARISON OF THE PERCENTAGE IDENTITY BETWEEN *Bam*HI-L OPEN READING FRAMES IN STRAINS AG876 AND B95-8

ORF	Function	Nucleotide homology (%)	Amino acid homology (%)
BLLF1	gp340/220	98.4	97.3
BLLF2	Early ORF	99.1	98.0
BLLF3	dUTPase	99.4	98.9
BLRF1	Putative membrane protein	100	100
BLRF2	Late ORF	99.2	100
BLRF3*	EBNA3a*	88.7	79.4

Note. Asterisk indicates that only the first exon of EBNA 3a (which is within *Bam*HI-L) has been considered and thus the percentage homology differs from that given in Sample *et al.* (1990). BLRF1 has recently been proposed to be a membrane protein (Barnett *et al.*, 1992).



Residue	B95-8	P3HR-1/AG876
17	H (His)	Q (Gln)
20	G (Gly)	R (Arg)
21	E (Glu)	D (Asp)
31	P (Pro)	L (Leu)
38	T (Thr)	A (Ala)
48	V (Val)	A (Ala)
60	Q (Gln)	K (Lys)
62	D (Asp)	N (Asn)
67	Q (Gln)	L (Leu)
115	V (Val)	I (Ile)
441	D (Asp)	A (Ala)
495	W (Trp)	R (Arg)
525	P (Pro)	S (Ser)
581	S (Ser)	P (Pro)
592	S (Ser)	I (Ile)
613	G (Gly)	S (Ser)
625	A (Ala)	T (Thr)
635	P (Pro)	S (Ser)
642	Q (Gln)	P (Pro)
672	N (Asn)	I (Ile)
673	P (Pro)	S (Ser)
710	I (Ile)	T (Thr)
754	Q (Gln)	K (Lys)
812	P (Pro)	T (Thr)

FIG. 5. Schematic map of B95-8 gp340 to indicate regions of functional significance. The polypeptide is represented by the horizontal bar and the repeat region is indicated. Amino acid position is shown along the bottom. The positions of substitutions between B95-8 and P3HR-I/AG876 are shown as vertical lines and are listed below. The region spliced out to give gp220 is indicated by the black bar. See text for sources of mapping immunologically significant domains.

that both type-A and type-B seem to have equal oncogenic potential in their association with Burkitt's lymphoma and nasopharyngeal carcinoma (Young *et al.*, 1987; Chen *et al.*, 1992). On the basis of the sequence homology and antibody binding data obtained in this study we conclude that a vaccine based on the B95-8 gp340/220 protein should be equally effective against both type-A and type-B strains of EBV.

ACKNOWLEDGMENTS

We are grateful to Drs. G. Bornkamm and J. Sample for gifts of plasmids, Drs. L. Qualtiere and D. Hayward for monoclonal antibodies, Dr. C. Rooney for cell line C2.BL16, and Nusrat Janjua for technical assistance. The contributions of Margaret Conway, Angela Evans, Renske Steenbergen, Emyr Thomas, and Susanna Vervoordeldonk during the preliminary stages of this project are appreciated. This work was supported by the Cancer Research Campaign.

REFERENCES

ADLDINGER, H. K., DELIUS, H., FREESE, O. K., CLARKE, J. and BORNKAMM, G. W. (1985). A putative transforming gene of Jijoye virus differs from that of Epstein–Barr virus prototypes. *Virology* **141**, 221–234.

ARRAND, J. R. (1992). Prospects for a vaccine against Epstein–Barr virus. *Cancer J.* **5**, 188–193.

ARRAND, J. R., YOUNG, L. S., and TUGWOOD, J. D. (1989). Two families of sequences in the small RNA-encoding region of Epstein–Barr virus (EBV) correlate with EBV types A and B. *J. Virol.* **63,** 983–986.

BAER, R., BANKIER, A. T., BIGGIN, M. D., DEININGER, P. L., FARRELL, P. J., GIBSON, T. J., HATFULL, G., HUDSON, G. S., SATCHWELL, S. C., SEGUIN, C., TUFFNELL, P. S., and BARRELL, B. G. (1984). DNA sequence and expression of the B95-8 Epstein–Barr virus genome. *Nature* **310,** 207–211.

BARNETT, B. C., DOLAN, A., TELFORD, E. A. R., DAVISON, A. J., and MCGEOCH, D. J. (1992). A novel Herpes simplex virus gene (UL49A) encodes a putative membrane protein with counterparts in other herpesviruses. *J. Gen. Virol.* **73,** 2167–2171.

BAUER, G. (1983). Quantitative analysis of the co-operative effect between inducers of Epstein–Barr virus antigen synthesis. *J. Gen. Virol.* **64,** 1337–1346.

BEISEL, C., TANNER, J., MATSUO, T., THORLEY-LAWSON, D., KEDZY, F., and KIEFF, E. (1985). Two major outer envelope glycoproteins of Epstein–Barr virus are encoded by the same gene. *J. Virol.* **54,** 665–674.

BORNKAMM, G. W., DELIUS, H., ZIMBER, U., HUDEWENTZ, J., and EPSTEIN, M. A. (1980). Comparison of Epstein–Barr virus strains of different origin by analysis of the viral DNAs. *J. Virol.* **35,** 603–618.

BORNKAMM, G. W., KNEBEL-DOEBERITZ, M. V., and LENOIR, G. M. (1984). No evidence for differences in the Epstein–Barr virus genome carried in Burkitt lymphoma cells and non-malignant lymphoblastoid cells from some patients. *Proc. Natl. Acad. Sci. USA* **81,** 4930–4934.

CHEN, X.-Y., PEPPER, S. DE V., and ARRAND, J. R. (1992). Prevalence of the A and B types of Epstein–Barr virus DNA in nasopharyngeal carcinoma biopsies from Southern China. *J. Gen. Virol.* **73,** 463–466.

DAMBAUGH, T., RAAB-TRAUB, N., HELLER, H., BEISEL, C., HUMMEL, M., CHEUNG, A., FENNEWALD, S., KING, W., and KIEFF, E. (1980). Variations among isolates of Epstein–Barr virus. *Ann. N.Y. Acad. Sci.* **354,** 309–325.

DAMBAUGH, T., HENNESSY, K., CHAMNANKIT, L., and KIEFF, E. (1984). U2 region of Epstein–Barr virus DNA may encode Epstein–Barr virus nuclear antigen 2. *Proc. Natl. Acad. Sci. USA* **81,** 7632–7636.

DE THÉ, G. (1985). Epidemiology of Epstein–Barr virus and associated diseases in man. *In* "The Herpesviruses" (B. Roizman, Ed.), *Vol. 3, pp. 25–103. Plenum, New York.*

DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395.

EMINI, E. A., SCHLEIF, W. A., SILBERKLANG, M., LEHMAN, D., and ELLIS, R. W. (1989). Vero cell-expressed Epstein–Barr virus (EBV) gp 350/220 protects marmosets from EBV challenge. *J. Med. Virol.* **27,** 120–123.

EPSTEIN, M. A., RANDLE, B. J., FINERTY, S., and KIRKWOOD, J. K. (1986). Not all potentially neutralizing vaccine-induced antibodies to Epstein–Barr virus ensure protection of susceptible experimental animals. *Clin. Exp. Immunol.* **63,** 485–490.

FINERTY, S., TARLTON, J., MACKETT, M., CONWAY, M., ARRAND, J. R., WATKINS, P. E., and MORGAN, M. J. (1992). Protective immunization against Epstein–Barr virus-induced disease in cottontop tamarins using the virus envelope glycoprotein gp340 produced from a bovine papillomavirus expression vector. *J. Gen. Virol.* **73,** 449–453.

GU, S., HUANG, T., MIAO, Y., RUAN, L., ZHAO, Y., HAN, C., XIAO, Y., ZHU, J., and WOLF, H. (1991). A preliminary study on the immunogenicity in rabbits and in human volunteers of a recombinant vaccinia virus expressing Epstein–Barr virus membrane antigen. *Chin. Med. Sci. J.* **6,** 241–243.

HARRIS, A., WALSH-ARRAND, J. E., CRAWFORD, D. H., LENOIR, G. H., and GRIFFIN, B. E. (1984). Molecular characterization of the Epstein–Barr virus DNA in three new Burkitt's lymphoma derived cell lines. *Mol. Biol. Med.* **2,** 135–150.

HELLER, M., DAMBAUGH, T., and KIEFF, E. (1981). Epstein–Barr Virus DNA IX. Variation among viral DNAs from producer and non-producer infected cells. *J. Virol.* **38,** 632–648.

HOFFMAN, G. J., LAZAROWITZ, G. S., and HAYWARD, S. D. (1980). Monoclonal antibody against a membrane antigen and a neutralizing antigen. *Proc. Natl. Acad. Sci. USA* **77,** 2979–2983.

JANOSSY, G., and AMLOT, P. (1987). Immunofluorescence and immunohistochemistry. *In* "Lymphocytes: A Practical Approach" (G. G. G. Klaus, Ed.), pp. 67–108. IRL Press, Oxford.

JENSON, H., FARRELL, P., and MILLER, G. (1987). Sequences of the Epstein–Barr virus (EBV) large internal repeat form the centre of a 16-kilobase pair palindrome of EBV (P3HR-1) heterogeneous DNA. *J. Virol.* **61,** 1495–1506.

JENSON, H., and MILLER, G. (1988). Polymorphisms of the region of the Epstein–Barr virus genome which disrupts latency. *Virology* **185,** 549–564.

KLEIN, G. (1989). Viral latency and transformation—The strategy of Epstein–Barr virus. *Cell* **58,** 5–8.

LEES, J. F., GOODEVE, A. C., ARRAND, J. E., GHOSH, A. K., JONES, P. H., and ARRAND, J. R. (1992). Detection of EBV DNA in post-nasal space biopsy tissue from asymptomatic EBV-seropositive individuals. *J. Med. Virol.* **37,** 30–38.

LUNG, M. L., CHANG, R. S., and JONES, J. H. (1988). Genetic polymorphism of natural Epstein–Barr virus isolates from infectious mononucleosis patients and healthy carriers. *J. Virol.* **62,** 3862–3866.

MADEJ, M., CONWAY, M. J., MORGAN, A. J., SWEET, J., WALLACE, L., QUALTIERE, L. F., ARRAND, J. R., and MACKETT, M. (1992). Purification and characterization of Epstein–Barr virus gp340/220 produced by a bovine papillomavirus expression vector system. *Vaccine* **10,** 777–782.

MORGAN, A. J., MACKETT, M., FINERTY, S., ARRAND, J. R., SCULLION, F. T., and EPSTEIN, M. A. (1988). Recombinant vaccinia virus expressing Epstein–Barr virus glycoprotein gp340 protects cottontop tamarins against EB virus-induced malignant lymphomas. *J. Med. Virol.* **25,** 189–195.

NEMEROW, G. R., MOLD, C., SCHWEND, V. K., TOLLEFSON, V., and COOPER, N. R. (1987). Identification of gp350 as the viral glycoprotein mediating attachment of Epstein–Barr virus (EBV) to the EBV/C3d receptor of B cells: Sequence homology of gp350 and C3 complement fragment C3d. *J. Virol.* **61,** 1416–1420.

PITHER, R. J., ZHANG, C. X., SHIELS, C., TARLTON, J., FINERTY, S., and MORGAN, A. J. (1992a). Mapping of B cell epitopes on the polypeptide chain of the Epstein–Barr virus major envelope glycoprotein and candidate vaccine molecule gp340. *J. Virol.* **66,** 1246–1251.

PITHER, R. J., NOLAN, L., TARLTON, J., WALFORD, J., and MORGAN, A. J. (1992b). Distribution of epitopes within the amino acid sequence of the Epstein–Barr virus major envelope glycoprotein, gp340, recognized by hyperimmune rabbit sera. *J. Gen. Virol.* **73,** 1409–1415.

POLACK, A., HARTL, G., ZIMBER, U., FREESE, U.-K., LAUX, G., TAKAKI, K., HOHN, B., GISSMANN, L., and BORNKAMM, G. W. (1984). A complete set of overlapping cosmid clones of M-ABA virus derived from nasopharyngeal carcinoma and its similarity to other Epstein–Barr virus isolates. *Gene* **27,** 279–288.

QUALTIERE, L. F., DECOTEAU, J. F., and NASR-EL-DIN, M. H. (1987). Epitope mapping of the major Epstein–Barr virus outer membrane glycoprotein gp350/220. *J. Gen. Virol.* **68,** 535–543.

ROWE, M., YOUNG, L. S., CADWALLADER, K., PETTI, L., KIEFF, E., and RICKINSON, A. B. (1989). Distinction between Epstein–Barr virus type A (EBNA 2A) and type B (EBNA 2B) isolates extends to the EBNA3 family of nuclear proteins. *J. Virol.* **63,** 1031–1039.

RYMO, L., LINDAHL, T., and ADAMS, A. (1979). Sites of sequence variability in Epstein–Barr virus from different sources. *Proc. Natl. Acad. Sci. USA* **76,** 2794–2798.

RYMO, L., LINDAHL, T., POVEY, S., and KLEIN, G. (1981). Analysis of restriction endonuclease fragments of intracellular Epstein–Barr

virus DNA and isoenzymes indicates a common origin of the Raji, NC-37, and F-265 human lymphoid lines. *Virology* **115,** 115–124.

SAIRENJI, T., BERTONI, G., MEDVECZKY, M. M., MEDVECZKY, P. G., NGUYEN, Q. V., and HUMPHREYS, R. E. (1988). Inhibition of Epstein–Barr virus (EBV) release from P3HR-1 and B95-8 cell lines by monoclonal antibodies to EBV membrane antigen gp350/220. *J. Virol.* **62,** 2614–2621.

SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

SAMPLE, J., YOUNG, L., MARTIN, B., CHATMAN, T., KIEFF, E., RICKINSON, A., and KIEFF, E. (1990). Epstein–Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B and EBNA-3C genes. *J. Virol.* **64,** 4084–4092.

SANGER, F., NICKLEN, S., and COULSON, A. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

SKEHEL, J. J., and WILEY, D. C. (1986). Antigenic variation in Hong Kong influenza virus haemagglutinins. *In* "Antigenic Variation in Infectious Diseases" (T. H. Birkbeck and C. W. Penn, Eds.), pp. 19–24. IRL Press, Oxford.

STADEN, R. (1987). Computer handling of DNA sequencing projects. *In* "Nucleic Acid and Protein Sequence Analysis: A Practical Approach" (M. J. Bishop and C. J. Rawlings, Eds.), pp. 173–217. IRL Press, Oxford.

SUGDEN, B. (1977). Comparison of Epstein–Barr viral DNAs in Burkitt lymphoma biopsy cells and in cells clonally transformed *in vitro*. *Proc. Natl. Acad. Sci. USA* **74,** 4651–4655.

TANNER, J., WEIS, J., FEARON, D., WHANG, Y., and KIEFF, E. (1987). Epstein–Barr virus gp 350/220 binding to the B-lymphocyte C3d receptor mediates adsorption, capping and endocytosis. *Cell* **50,** 203–213.

TANNER, J., WHANG, Y., SAMPLE, J., SEARS, A., and KIEFF, E. (1988). Soluble gp350/220 and deletion mutant glycoproteins block Epstein–Barr virus adsorption to lymphocytes. *J. Virol.* **62,** 4452–4464.

THOMAS, J. A., and CRAWFORD, D. H. (1989). Epstein–Barr virus-associated B-cell lymphomas in AIDS and after organ transplantation. *Lancet* **i,** 1075–1076.

THORLEY-LAWSON, D. A., and POODRY, C. A. (1982). Identification and isolation of the main component (gp350/gp220) of Epstein–Barr virus responsible for generating neutralizing antibodies *in vivo*. *J. Virol.* **43,** 730–736.

WALLACE, L. E., WRIGHT, J., ULEATO, D. O., MORGAN, A. J., and RICKINSON, A. B. (1991). Identification of two T-cell epitopes on the candidate Epstein–Barr virus vaccine glycoprotein recognized by CD4+ T-cell clones. *J. Virol.* **65,** 3821–3828.

WU, T. C., MANN, R. B., CHARACHE, P., HAYWARD, S. D., STAAL, S., LAMBE, B. C., and AMBINDER, R. F. (1990). Detection of EBV gene expression in Reed-Sternberg cells of Hodgkin's disease. *Int. J. Cancer* **46,** 801–804.

YOUNG, L. S., YAO, Q. Y., ROONEY, C. M., SCULLEY, T. B., MOSS, D. J., RUPANI, H., LAUX, G., BORNKAMM, G. W., and RICKINSON, A. B. (1987). New type B isolates of Epstein–Barr virus from Burkitt's lymphoma and from normal individuals in endemic areas. *J. Gen. Virol.* **68,** 2853–2862.

ZHANG, P.-F., KLUTCH, M., ARMSTRONG, G., QUALTIERE, L., PEARSON, G., and MARCUS-SEKURA, C. J. (1991). Mapping of the epitopes of Epstein–Barr virus gp350 using monoclonal antibodies and recombinant proteins expressed in *Escherichia coli* defines three antigenic determinants. *J. Gen. Virol.* **72,** 2747–2755.

ZIMBER, V., ADLDINGER, H. K., LENOIR, G. M., VUILLAUME, M., KNEBEL-DOEBERITZ, M. V., LAUX, G., DESGRANGES, C., WITTMANN, P., FREESE, U. K., SCHNEIDER, V., and BORNKAMM, G. W. (1986). Geographical prevalence of two types of Epstein–Barr virus. *Virology* **154,** 56–66.