Volume 8 Number 19 1980 Nucleic Acids Research

The baboon endogenous virus genome. II. Provirus sequence variations in baboon cell DNA

Maurice Cohen*†, Norman Davidson†, Raymond V.Gilden*, Robert M.McAllister+, Margery O. Nicolson+ and Robert M.Stephens*

*Biological Carcinogenesis Program, Frederick Cancer Research Center, Frederick, MD 21701, †Department of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, and +Childrens Hospital of Los Angeles, Los Angeles, CA 90054, USA

Received 11 July 1980

ABSTRACT

Restriction analysis of the approximately 100 integrated baboon endogenous virus (BaEV) proviruses in baboon cells and tissues has revealed two major sequence variations, both in the gag gene region of the genome. One, a 150 nucleotide pair insert, is present in a small proportion of the proviral DNAs of some baboons, but is present in the majority of the proviral DNAs of other baboons. The second, a Bam HI recognition sequence located 2.25 kb from the proviral 5' end, is missing or modified in approximately one-half of the integrated genomes. We consider the possibility that accumulation of proviruses not containing the 0.15 kb insert is correlated with viral activation and expression since it is this form that is a replication intermediate in freshly infected permissive cells. It is evident from these initial studies that the organization of the multiple BaEV proviruses in baboon DNA has undergone modification during evolution.

INTRODUCTION

Evidence is accumulating which indicates that retrovirus DNA sequences constitute a significant portion of the mammalian genome. Many mouse strains, for instance, contain 1000-1800 chromosomal copies of intracisternal type A particle-specific DNA representing 0.2-0.35% of the mouse genome (1,2). In primates six different endogenous retroviruses have been detected and characterized. These are owl monkey type C and squirrel monkey type D viruses (New World retroviruses) and baboon, macaque, and colobus type C viruses and langur monkey type D virus (Old World retroviruses). In each case the respective retrovirus is integrated in the host genome in multiple copies (3).

In baboon, multiple copies of three different primate retroviruses have been detected thus far. Baboon endogenous virus (BaEV) sequences are reiterated approximately 100 times (4). The macaque type C retroviruses, MAC-1 and MMC-1 (5,6), which are closely related to each other but which have limited sequence homology with BaEV, are also highly reiterated in baboon cells (M. Tainsky, manuscript submitted). Further, the type D retrovirus of the

spectacled langur, which is unrelated to BaEV and to the macaque viruses (5), is homologous to sequences present in baboon cell DNA (7). Together these proviruses consititute 0.02-0.04% of the baboon genome. The significance of these multiple genomes is totally unclear; speculations range from a role in development or a stage in the life cycle of an obligate parasite (8) to no function whatsoever (9,10).

To aid in understanding the complexities of integrated retroviral genomes, we are analyzing the organization of baboon endogenous virus sequences in both virus producing and nonproducing baboon cells. We have chosen to study these sequences in DNA of BEF-3, a nonmalignant virus-productive baboon embryo fibroblast cell strain that contains the same number (~100) of BaEV proviral genome equivalents as several other baboon cells (4), DNA of RD(BAB8-K), a human cell line productively infected with BaEV, and DNA extracted from tissues of four apparently healthy baboon juveniles.

Since BaEV is seldom expressed in baboon cells and since baboon DNA is not infectious when tested on permissive cells (4), it is probable that most BaEV proviruses in baboon DNA are defective or are under negative host control (11). BEF-3 DNA, in contrast, is expected to contain some nondefective BaEV genomes because BEF-3 cells are virus-productive and its DNA is infectious. Thus, we were interested in examining the heterogeneity of provirus sequences within BEF-3 DNA and comparing it to the BaEV sequences in DNA of baboon tissues.

One indication of the BaEV provirus heterogeneity in baboon DNA was the observation that molecular hybrids formed between BaEV complementary DNA (cDNA) and baboon cell DNA have a lower thermal stability and a broader melting profile than comparable hybrids formed between this cDNA and RD(BAB8-K) DNA which contains 5-6 BaEV proviral copies (N. Rice, unpublished results).

In an initial study we reported a detailed restriction enzyme map of the BaEV unintegrated linear DNA isolated from freshly infected permissive human cells (12). This map was essential for interpretation of the complex proviral DNA fragment patterns observed in cellular DNA digests. Earlier studies reported that the multiple integrated BaEV viral genomes are not tandemly linked in either BEF-3 or RD(BAB8-K) (13).

In this report we show that the BaEV provirus heterogeneity in BEF-3 DNA includes an alteration of a Bam HI recognition sequence in approximately one-half of the provirus genomes and the presence of a specific 150 nucleotide pair deletion in most but not all proviruses. Further, we report that the ratio of deleted to undeleted provirus genomes is significantly different in

the DNA of these baboons from that in the DNA of another baboon and BEF-3 cells.

MATERIALS AND METHODS

Material. Restriction endonucleases were purchased from Bethesda Research Laboratories (Bethesda, MD) and Seakem agarose from Marine Colloids, Inc. (Rockland, ME). α-^{32}P deoxyribonucleotides (specific activity = 300-400 Ci/mMol) were obtained from Amersham-Searle (Arlington Heights, IL). The vertical electrophoresis gel apparatus was purchased from Watson Products, Inc. (Altadena, CA).

Cells and tissues. The baboon embryo fibroblast cell strain, BEF-3, and the baboon tissues were obtained from Dr. R. Heberling (4). The RD(BAB8-K) producer cell line was derived by infection of human rhabdomyosarcoma (RD) cells (14) with the BAB8-K isolate of BaEV (20).

DNA isolation. High molecular weight cellular and tissue DNAs were isolated by lysing washed cells or homogenized tissues in medium containing 0.01 M Tris-HCl pH 8.1, 0.1 M NaCl, 0.1 M EDTA, 100 µg/ml protinase K, and 0.5% SDS at 37°C for 18 hr. NaCl was then added bringing its concentration to 0.2 M. DNAs were gently mixed with an equal volume of chloroform:isoamyl alcohol (24:1) and after centrifugation, the aqueous phases were removed; the chloroform extractions were repeated. DNAs were precipitated with ethanol, spooled on glass rods, and dissolved in buffer containing 0.01 M Tris-HCl-NaCl EDTA, pH 7.4. They were then sequentially treated with RNase A (100 µg/ml, 37°C, 18 hr) and protinase K (50 µg/ml, 37°C, 18 hr). Sodium chloride was added to a concentration of 0.2 M and the DNAs were extracted twice with chloroform-isoamyl alcohol as before. Following ethanol precipitation and spooling of the DNAs, they were banded by isopycnic centrifugation in cesium chloride (starting density of 1.68 g/cc) in a SW50.1 rotor at 45,000 rpm for 50 hr.

Blotting, cDNA synthesis, RNA isolation. Our procedures for Southern elution (15), viral RNA isolation, and cDNA synthesis have been previously reported (12).

DNA elution from agarose. To locate restriction fragments of a particular length in the preparative 1% agarose gel, pBR322 DNA fragments of a known length (16) were electrophoresed in the outside lanes and located by ethidium bromide staining. This allowed us to cut out regions of the gel corresponding to the desired DNA lengths. The agarose was dissolved in a saturated solution of potassium iodide containing 0.005 M sodium phosphate pH 7.8, and 0.001 M

β-mercaptoethanol. The mixture was incubated at 50°C for 1 hr and then at 20°C for 18 hr. The solution was passed over a column of hydroxyapatite equilibrated with 0.01 M sodium phosphate at 55°C. Two aliquots of the same buffer were used to rinse the column. DNA was eluted from the hydroxyapatite with 0.5 M sodium phosphate at 55°C and was dialyzed extensively against low salt buffer to remove the phosphate. DNA was concentrated by extraction of water with n-butanol.

RESULTS

Proviral Bam HI fragments. As in our previous report on unintegrated BaEV DNA (13), we have characterized the integrated BaEV proviruses extensively with respect to the restriction enzyme Bam HI. The Bam HI fragments of integrated BEF-3, RD(BAB8-K), and unintegrated BaEV DNA that are detected in Southern blots (14) by a [^{32}P]cDNA probe representative of the BaEV genome (Fig. 1) are summarized in Table 1. Fragments of low or variable autoradiographic intensity are so designated. The faint bands of 2.5 kb, 2.3 kb, and 1.45 kb that are occasionally observed in Bam HI digests of BEF-3 DNA (Fig. 2) are due to ribosomal contamination of the viral probe, as demonstrated in separate blots by hybridization with ^{32}P-labeled ribosomal cDNA (data not shown). We have previously observed that Bam HI digests of unintegrated

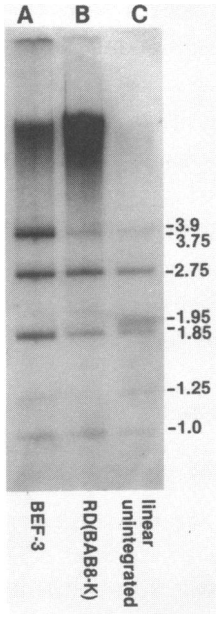

Figure 1. Comparison of integrated and unintegrated BaEV provirus Bam HI fragments. Cellular DNAs isolated from the cells indicated below lanes A and B and from BaEV unintegrated linear DNA isolated from freshly infected human cells (lane C) were digested with Bam HI, electrophoresed in a 1% agarose gel for 5 hr at 4.2 V/cm, transferred to a 0.45μ nitrocellulose sheet and hybridized to BaEV [^{32}P]cDNA as described in experimental procedures. The hybrids were detected by autoradiography. Restriction digests of pM2 and pBR322 DNAs were electrophoresed in adjacent lanes of the gel as length standards. Fragment lengths are given in kilobase pairs.

Table 1. The Bam HI restriction fragments of BaEV in digests of BEF-3 DNA, RD(BAB8-K) DNA, and linear unintegrated DNA detected with a BaEV [^{32}P]cDNA probe. Designation of the fragments (in parentheses) refers to the restriction map in Fig. 3. In addition to these bands, ribosomal DNA bands of lengths 2.5, 2.3, and 1.45 kb were occasionally observed in BEF-3 DNA digests.

Band (kb)	BEF-3	RD(BAB8-K)	Linear Unintegrated
3.9	faint (B+C)	nd	nd
3.75	strong (B+C)	faint	faint (D+E)
2.75	strong (C)	strong (C)	strong (C)
1.95	nd	nd	strong (E)
1.85	strong (D)	strong (D)	strong (D)
1.25	strong (A)	v. faint (A)	strong (A)
1.18	faint, var. (B)	nd	nd
1.0	strong (B)	strong (B)	strong (B)

nd = not detected; var = variable; v = very.

linear BaEV DNA give a fragment of 3.8 kb (12). This, however, was shown to be due to incomplete cleavage at Bam HI site (DE) (Fig. 3) and is not related to the 3.75 kb Bam HI fragment of BEF-3 DNA which, as we will show below, consists of fragments B and C.

The three intense bands of 2.75 kb, 1.85 kb, and 1.0 kb are expected in digests of integrated proviral DNA digests (Figs. 1 and 2) because they come

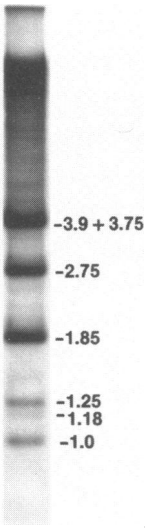

Figure 2. Bam HI proviral fragments of BEF-3 DNA. As described in the legend to Fig. 1; the long autoradiographic exposure which was necessary to reveal minor bands resulted in overexposure and incomplete resolution of the 3.9-3.75 kb fragments.

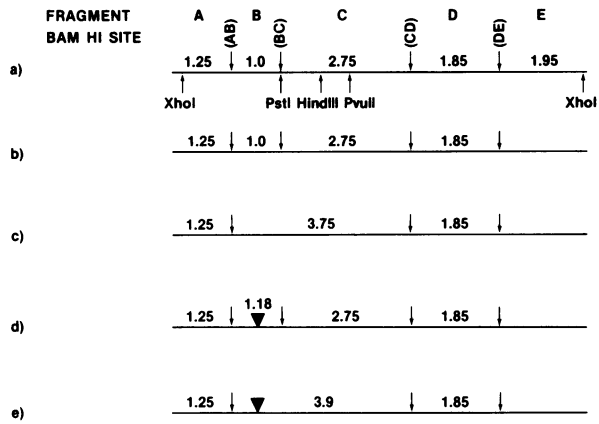

Figure 3. Restriction map of unintegrated and integrated BaEV proviruses. The restriction enzyme sites for BaEV unintegrated linear DNA (a) were previously described (Cohen et al., 1980). The Bam HI cleavage sites [(AB), (BC), (CD), (DE)] and the fragments generated (A-E) are given above the map. The Bam HI cleavage sites and the fragments observed in digests of baboon chromosomal DNAs are shown in b-e. ▼denotes a 150 nucleotide pair insertion with respect to the unintegrated DNA form that is observed in B fragments or B + C fusion fragments. Although the 1.25 kb Bam HI fragment is denoted at the 5' end of integrated BaEV proviruses (b-e), it is uncertain whether the 5' distal Bam HI site resides within the genome or in the baboon flanking sequence (see text). The restriction sites observed in the unintegrated DNA (a) are present in many, but probably not all integrated proviruses (b-e) as well.

from internal Bam HI fragments (Fig. 3) which are obviously present in many of 90-100 endogenous genomes of BEF-3 and 5-6 endogenous genomes of RD(BAB8-K).

The four bands in BEF-3 digests which are not expected from the pattern of the unintegrated linear viral DNA and which are of particular interest are: 3.9 kb (faint and poorly resolved), 3.75 kb (strong), 1.25 kb, and 1.18 kb (faint but variable intensity). We discuss these various bands below.

As shown in Fig. 3a, the two end fragments of the unintegrated linear BaEV DNA have lengths of 1.25 kb and 1.95 kb. If, as is the case for most retroviruses, there are many diverse integration sites with different flanking sequences (17,18), each end fragment would be contained in a unique fragment of length greater than 1.25 kb or 1.95 kb, respectively, depending upon the fortuitous position of the next Bam HI site in the host flanking sequence. Because of the high copy number, one would not expect a clear resolution of the single copy bands containing these fragments. In agreement with this prediction, a discrete band of 1.95 kb is not seen in digests of either chromo-

somal DNA (Fig. 1). However, unexpectedly, a band corresponding to approximately 1.25 kb is generated by Bam HI digestion of BEF-3 DNA (Figs. 1 and 2). In addition, a very faint band of this length is sometimes seen in the corresponding digest of DNA from the infected human cell RD(BAB8-K) (Fig. 1).

5' end Bam HI fragment. In order to determine whether these 1.25 kb fragments correspond to the 1.25 kb 5' end fragment of unintegrated linear BaEV DNA, the two cellular DNAs and the unintegrated viral DNA were digested sequentially with Bam HI and Xho I before electrophoresis, blotting, and hybridization. The latter enzyme cleaves the unintegrated linear BaEV DNA only in its terminal repetition (0.33 kb from the 5' end and 0.21 kb from the 3' end) (12). If the 1.25 kb Bam HI band of the two cellular DNAs represents the 5' ends of integrated BaEV genomes, Xho I cleavage should result in the disappearence of the band with the concomitant appearance of a 0.9 kb fragment. As shown in Fig. 4, Xho I did cleave the 1.25 kb Bam HI fragment resulting in the appearance of a new 0.9 kb band. (Also seen in the Xho I/Bam HI double digest is a 1.74 kb band, the expected product of Xho I cleavage near the 3' end of unintegrated linear and integrated proviral genomes.) Thus, the 1.25 kb fragment generated by Bam HI cleavage of integrated BaEV proviruses does correspond to the 1.25 kb 5' end fragment of unintegrated linear BaEV DNA, at least as far distal as the Xho I site, 0.33 kb from the 5' end. Furthermore, this experiment strongly suggests that the terminal repetition

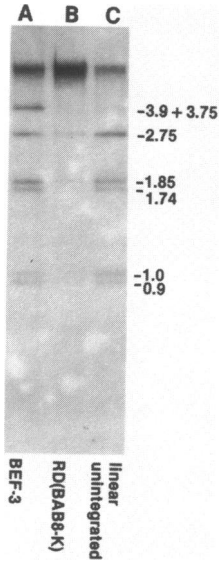

Figure 4. Internal provirus restriction fragments of integrated and unintegrated BaEV genomes. As described in the legend to Fig. 1 except that the DNAs were digested sequentially with Bam HI and Xho I and were electrophoresed in a 1.5% agarose gel for 5 hr at 4.7 V/cm.

present in the unintegrated BaEV DNA is also present in the integrated BaEV proviruses and that the proviral internal restriction fragments are the same lengths as those of the unintegrated DNA.

As was previously noted in Bam HI digests of BaEV unintegrated linear DNA, the internal 1.0 kb fragment often appears as a more intense band than does the 5' end 1.25 kb fragment (12). The reason for this is not known, but a possible explanation is that the cDNA probe may not be truly representative. As shown in Figs. 1 and 2, digests of BEF-3 DNA also result in a "reverse" intensity of these two Bam HI fragments.

We conclude that at least some of the integrated BaEV proviral genomes are cleaved near their 5' end releasing a 1.25 kb fragment. Possible explanations for this unexpected result are discussed later.

Origin of 3.75 and 3.9 kb Bam HI fragments. We now wish to consider the strong 3.75 kb and the faint 3.9 kb and 1.18 kb bands of BEF-3 DNA (Table 1). We shall present evidence that the 3.75 kb fragment is the sum of the 1.0 kb and 2.75 kb bands (Fig. 3, fragments B and C) indicating that the Bam HI cleavage site marked (BC) in Fig. 3 is missing or modified so as to be resistant to digestion in many of the integrated genomes (Fig. 3c). We will then present evidence that the 3.9 kb band (Fig. 3e) is related to the 3.75 kb band, except that there is an insertion of an additional sequence of length 0.15 kb within the fragment B. The 1.18 kb fragment (Fig. 3d), thus, derives from a few proviral genomes which contain the insertion in fragment B, but which have retained Bam HI site (BC). For simplicity, we will refer to the 1.18 kb and 3.9 kb Bam HI fragments as coming from "inserted" genomes and the 1.0 kb and 3.75 kb fragments as coming from "uninserted" genomes. This does not presuppose the actual ancestral relationship between the two forms and the 3.9 kb fragment may, in fact, antedate the 3.75 kb form.

The intensity of the 3.75 kb BEF-3 DNA Bam HI fragment suggested that it was present in molar mass ratio approximately equal to that of the other four major Bam HI fragments (Figs. 1 and 2). However, the presence of an internal fragment of this length was inconsistent with the unintegrated DNA restriction map (12). The origin of this fragment was determined from chromosomal blots of BEF-3 DNA digested with both Bam HI and Pst I. Pst I cleaves the unintegrated linear BaEV DNA once at a site about 2.25 kb from the 5' end of the genome, a position very close to Bam HI cleavage site (BC) in Fig. 3 (12). In the double digest with Pst I, the 3.75 kb Bam HI fragment disappeared (Fig. 5) indicating that it derived from fusion of the 1.0 and 2.75 kb Bam HI fragments. The 3.9 kb fragment is also totally or almost totally digested and

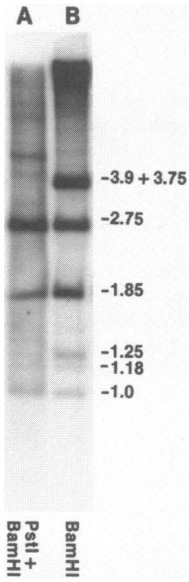

Figure 5. Origin of the BEF-3 DNA 3.75 kb Bam HI fragment. As described in the legend to Fig. 1 except that BEF-3 DNA was digested with either Bam HI plus Pst I (lane A) or with Bam HI alone (lane B).

the significance of this will be considered below.

In order to confirm this interpretation and to obtain reliable information about the nature of the faint 3.9 kb band, we isolated the 3.75 kb and 3.9 kb fragments by agarose gel electrophoretic fractionation and studied their digestion with other restriction enzymes. Two fractions of DNA were recovered from the preparative gel: a primary slice corresponding to DNAs of 3.75-3.9 kb in length and a pool fraction consisting of DNA recovered from the two combined gel slices adjacent on both sides to the primary slice.

As shown in Fig. 6 (lane C), the primary fraction consists principally of the 3.75 kb and 3.9 kb fragments and is somewhat enriched in the latter. The pooled fractions from the two sides of the primary fraction (Fig. 6, lane B) are depleted in the 3.9 kb fragment but contain a fair amount of the 3.75 kb fragment and other higher and lower molecular weight components including a trace of the 2.75 kb fragments.

To ensure that the 3.75 kb and 3.9 kb DNA bands were not the result of incomplete digestion with Bam HI, we redigested the purified DNAs with Bam HI and found no change in the pattern of hybridization (data not shown); however, when the DNAs were digested with restriction enzymes known to cleave within the unintegrated DNA 2.75 kb Bam HI fragment (12), both bands disappeared and two intensely and one faintly hybridizing bands appeared.

With each enzyme tested, the appearance of a faint band was most pro-

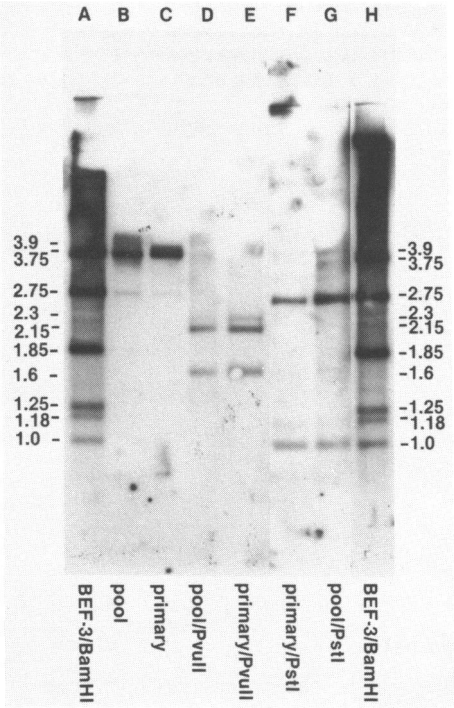

Figure 6. Cleavage of purified 3.75 kb and 3.9 kb Bam HI fragments. Bam HI cleavage products of BEF-3 DNA were isolated from an agarose gel following electrophoresis as described in Materials and Methods. The pool fraction DNA (B) was enriched in 3.75 kb molecules. The primary fraction DNA (C) contained both 3.75 kb and 3.9 kb molecules. Lanes D and E contained Pvu II cleavage products of the pool and primary fraction DNAs, respectively. Lanes F and G contained Pst I cleavage products of primary and pool fraction DNAs, respectively. Lanes A and H contained BEF-3 DNA digested with Bam HI. Conditions of Southern transfer and hybridization are as described in the legend to Fig. 1.

nounced in the digest of DNA containing the 3.9 kb band. That is, when DNA from the primary fraction (Fig. 6, lane C) was digested with Pvu II, intense bands of 2.15 and 1.6 kb and a less intense band of 2.3 kb were seen (Fig. 6, lane E). The intense bands are those expected from Pvu II cleavage of a 3.75 kb fragment, if that fragment originated from a fusion of the 1.0 and 2.75 kb Bam HI fragments. Pool fraction DNA (lacking the 3.9 kb band) that was cleaved with Pvu II gave rise to the same two prominent bands and to only a trace of the 2.3 kb fragment (Fig. 6, lane D). The straightforward interpretation of these results is that the less intense 3.9 kb fragment represents a subset of 3.75 kb molecules, but with an additional 0.15 kb of sequence at some point to the left of the Pvu II site (see Fig. 3). The intense 1.6 kb fragment is a product common to digestion of both purified DNAs, the intense 2.15 kb fragment is the result of cleavage of 3.75 kb molecules, and the less intense 2.3 kb fragment is due to cleavage of the less intense 3.9 kb molecules.

Similarly, when the two isolated DNAs were digested with Pst I, electrophoresed and blotted, DNA from the primary slice gave two intense bands (approximately 2.75 and 1.0 kb) and one faint band (1.18 kb) (Fig. 6, lane F),

while DNA from the pool fraction gave the same two intense bands but only a trace of the faint 1.18 kb fragment (Fig. 6, lane G). This is consistent with the previous interpretation that the 3.9 kb fragment is a subset of 3.75 kb molecules that contains an additional 150 nucleotide pairs. These data allowed us to map the position of the additional DNA in the 3.9 kb fragment to the left of the Pst I site.

The 1.18 kb band. The remaining unexplained band in Bam HI digests of BEF-3 DNA is also a fragment of 1.18 kb. Because the unique Pst I site in BaEV unintegrated linear DNA is closely adjacent to Bam HI site (BC) (12; see Fig. 3), this 1.18 kb fragment most probably originated in an analogous fashion to the faint 1.18 kb fragment which is released from 3.9 kb molecules by Pst I digestion. That is, some BEF-3 provirus genomes contain both Bam HI site (BC) and the 0.15 kb inserted sequence (Fig. 3d). When cleaved by Bam HI, these genomes release fragments of lengths 1.18 kb, 2.75 kb, 1.85 kb, and possibly 1.25 kb (see above).

Baboon tissue DNA digests. Having arrived at an interpretation of the restriction fragments and their heterogeneity in BEF-3, we examined the DNA of various baboon tissues in order to compare the heterogeneity in the restriction sites in these various animals with that of the cell strain BEF-3. DNA from six tissues of four apparently healthy Papio cynocephalus juveniles was analyzed by Bam HI digestion (Fig. 7). Five of the six baboon tissue DNAs gave a Bam HI fragment pattern (lanes A-D, F) different in some respects from that seen in parallel digests of DNA isolated from three different passages of BEF-3 cells (lanes G-I). The digests of these tissue DNAs revealed a reverse intensity band pattern to that of BEF-3 DNA (in which the 3.75 kb band is stronger than the 3.9 kb band and the 1.0 kb band is stronger than the 1.18 kb band). It is striking that the DNA from a fourth baboon (designated animal 658) gave a Bam HI pattern identical to that of BEF-3 and unlike that from the other baboon tissues (lane E).

To further clarify the differences between the DNAs of animal 658 from those in other baboons, DNAs from animal 658, animal 984 (spleen) and BEF-3 cells were digested sequentially with Pvu II and Bam HI (Fig. 8). As in the earlier Pvu II digestion of isolated BEF-3 Bam HI fragments of 3.9 kb and 3.75 kb (Fig. 6), the Bam HI fragments from these three DNAs released 2.3 and 2.15 kb fragments in the double digest in proportion to the ratio, respectively, of 3.9 kb and 3.75 kb fragments present in the Bam HI digest alone. Thus, in the Bam HI digest of DNA from animal 984 (lane A), the 3.9 kb band was more intense than the 3.75 kb band, while in the double digest with Pvu II (lane B),

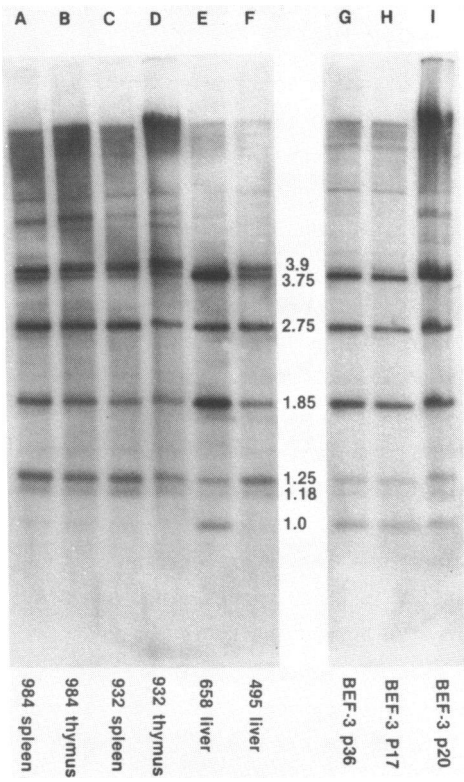

Figure 7. Bam HI cleavage products of baboon tissue and cell DNAs. DNAs extracted from six tissues of four baboons (lanes A-F) and from three passage levels of BEF-3 cells (lanes G-I) were digested with Bam HI. Details of electrophoresis, etc. are, as described in the legend to Fig. 1. The animal number and source of tissue, or passage number of BEF-3 cells is given under the respective lane.

the 2.3 kb band was more intense than the 2.15 kb band. On the other hand, BEF-3 DNA and animal 658 DNA which had 3.9 kb bands of lower intensity than 3.75 kb bands in Bam HI digests (lanes E and F), had 2.3 kb bands of lower intensity than 2.15 kb bands in the double digest (lanes C and D). These results support our contention that the 3.75 kb and 3.9 kb Bam HI fragments of BEF-3 are variants of the same fragment and show that this is also true of these fragments in DNA of baboon tissues.

Pvu II completely cleaved the 3.75 kb and 3.9 kb Bam HI fragments in the double digests (Fig. 8, lanes B-D) as expected from the position of the Pvu II site in the BaEV unintegrated linear DNA Bam HI fragment C (see Fig. 3), but unexpectedly failed to completely cleave the 2.75 kb fragments in the same digests. The 2.75 kb Bam HI fragments, however, were completely cleaved by Hind III (Fig. 8, lanes G and H), an enzyme that has its single restriction site in the BaEV unintegrated linear DNA also in Bam HI fragment C (Fig. 3). This result indicates that the failure of Pvu II to completely cut 2.75 kb

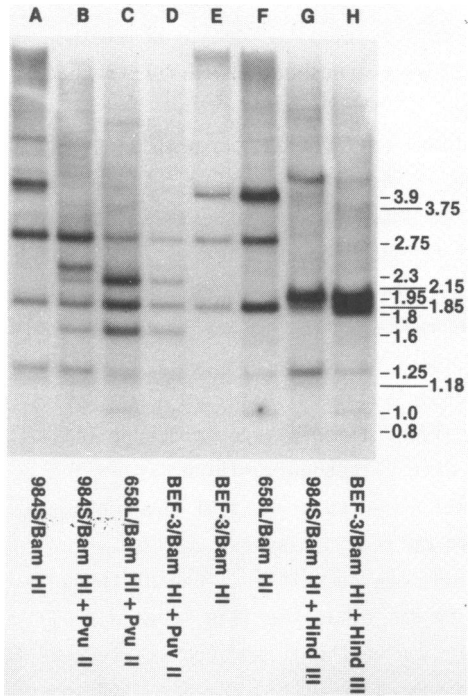

Figure 8. Digests of baboon tissue and cell DNAs. As shown below the figure, DNA from the spleen of animal 984 (984S), the liver of animal 658 (658L), and BEF-3 cells was digested with the indicated enzyme or enzymes. Conditions are as given in the legend to Fig. 1.

Bam HI fragments is not due to the presence of 2.75 kb Bam HI fragments that derive from another part of the BaEV genome. It is unclear why Pvu II did not completely cleave the Bam HI 2.75 kb fragment, but it may indicate the presence of another provirus heterogeneity in the Pvu II recognition sequence. If this is so, the occurrence of the Pvu II site heterogeneity is correlated with the heterogeneity of Bam HI cleavage at site (BC) (Fig. 3).

DISCUSSION

We have taken advantage of the high copy number of endogenous baboon proviruses to characterize their genome organization in Southern blots of chromosomal DNA. Restriction fragments cleaved from within integrated genomes result in intense bands whereas faint bands are produced by provirus end fragments that are attached to diverse host sequences. There may be other BaEV-like sequences present in the baboon genome which are detected by hybridization but which are sufficiently divergent so that their restriction site distribution is quite different from that of a competent BaEV provirus genome. These sequences would be effectively single copy and would contribute to the background of labeling seen in Southern blot autoradiograms. However, our

observations show that there are a substantial number of genomes with conserved restriction endonuclease sites which give rise to intense bands in digests and some of these bands are identical to those observed for the unintegrated linear viral DNA.

Infectious baboon DNAs could be isolated only from virus-producing baboon cell cultures and the infectivities of such DNAs for permissive cells was substantially less (1/2 to 1/100) than that of DNAs from exogenously infected dog or human producer cells containing only 5-6 integrated proviruses (4). Thus, we anticipated that the BaEV genomes present in baboon DNA were either mostly incomplete or else situated in the baboon genome in such a way that they could not be expressed, or both. Nevertheless, BaEV has been repeatedly isolated from baboon tissue by cocultivation with permissive cell strains (19), by spontaneous release from cultured baboon placental or embryonic tissue (4), and by treatment of baboon cell cultures with 5-iododeoxyuridine (20). The experiments described in this report detail observed aspects of the considerable BaEV provirus heterogeneity in baboon cells and tissues.

There are two particularly interesting heterogeneities. One of these is heterogeneity in sensitivity to cleavage by Bam HI at the site labeled (BC) in Fig. 3. The other is that a small fraction of the integrated proviruses of BEF-3 have an additional 0.15 kb sequence somewhere within the Bam HI fragment B (Fig. 3). Both heterogeneities appear to map within the BaEV gag gene region (21).

One possibility for the observed heterogeneity in Bam HI site (BC) is that a base in the Bam HI recognition sequence, 5'-GGATCC-3', may be methylated, thus altering its ability to be cleaved by this enzyme. Even though it was reported that DNA is not protected from Bam HI cleavage when either the terminal cytosine is methylated (22) or when the adenine is methylated (23), methylation of the penultimate cytosine residue renders the sequence uncleavable by Bam HI (24). Since 5-methyl cystosine occurs almost exclusively in mammalian cells in the dinucleotide, CpG (25), if the Bam HI sequence remains intact, either there is an exception to this rule in many of the endogenous BaEV genomes such that the penultimate cytosine is methylated or else methylation does not account for the observed heterogeneity.

One recombinant DNA clone that we have selected and characterized from a library of BEF-3 DNA fragments that were ligated into the λ vector, Charon 4A (Cohen et al., manuscript in preparation), is not cleaved by Bam HI at site (BC) (Fig. 3) and consequently releases a 3.75 kb fragment. Since it would be unexpected for a prokaryote (E. coli) to methylate a DNA in the identical

sites it is methylated in its original eukaryotic host, we conclude that at least some proviruses in BEF-3 DNA are intransigent to cleavage by Bam HI at site (BC) not because they are methylated, but because they contain an altered DNA sequence.

While only a small fraction of the integrated proviruses of BEF-3 have the additional 0.15 kb sequence in fragment B, more than one-half of the integrated proviruses of five baboon tissues contain the additional sequence (Fig. 7). Since BEF-3 is virus productive and its DNA is infectious whereas most baboon tissues (including those from baboons 984 and 932) are not virus productive and their DNAs are not infectious, this suggests the hypothesis that the genome or genomes which are being expressed to produce virus do not contain the additional 0.15 kb sequence. In addition, most BaEV proviruses are probably incapable of expression because of other defects so that accumulation of uninserted proviruses in an individual should increase the likelihood of having a genome that is completely nondefective. In support of this hypothesis, we have found that the unintegrated linear DNAs purified from permissive cells infected with three different isolates of BaEV [(BEF-3, M7, and RD(BAB8-K)] contain detectable levels of only the uninserted 1.0 kb form of Bam HI fragment B (Cohen et al., unpublished results).

DNA from the liver of one animal, baboon 658, was similar to BEF-3 DNA in that most of its genomes are of the uninserted lengths 3.75 kb and 1.0 kb. It is not known whether baboon 658 cells were releasing virus; the infectivity of its DNA is currently being tested.

The questions of whether the 0.15 kb DNA sequence was inserted into proviruses of the standard 8.8 kb length (12) or deleted from proviruses of 8.95 kb length, and when and how this event occurred may be impossible to answer. Since the two forms were first noted in the BEF-3 cell strain which is virus productive, the possibility was considered that this heterogeneity arose as the result of DNA changes which occurred after prolonged in vitro culture. However, DNA isolated from BEF-3 cells which had been in culture for varying times revealed identical Bam HI fragment patterns (Fig. 7, lanes G-I) suggesting that further observable changes were no longer occurring.

The fact that the baboon DNAs which we have studied fall into two basic patterns with respect to the 0.15 kb sequence is thus far unexplained. One possibility is that the DNAs were isolated from baboons of two different Papio cynocephalus group subspecies, each having a distinctive level of proviruses with and without inserts. The criteria for distinguishing between the subspecies (coat color, geographic origin) were followed in classification of

these animals (R. Heberling, personal communication); but, as the individuals were all captured in the wild (except for one animal born in captivity) at different times and locations, we have no information regarding their possible familial relationships.

On the other hand, we may be observing the results of further provirus integrations which have occurred during development of the individual animal. Since BaEV is not normally expressed in baboon tissue (20), there should be no interference barrier to superinfection. Thus, virus activation could lead to further viral genome integration. BaEV is known to penetrate baboon fibroblasts and subsequently to express its gene products (26). Such a possibility might explain the acquisition of many genomes of the uninserted 3.75 kb type in animal 658 and BEF-3, but it would not explain the loss (relative to the other tissue DNAs) of genomes of the inserted 3.9 kb type in these same two DNAs. Therefore, the data are not completely explained by this hypothesis.

It is interesting to note that this is not the first reported case of heterogeneity in the gag region of a mammalian proviral genome. The exogenously acquired mouse mammary tumor virus (MMTV) proviruses in mammary tumors can be distinguished from the endogenous MMTV genomes of BALB/c mice since the former contain a Pst I restriction site 1.5 x 10^6 daltons (2.3 kb) from the 5' end of the genome that is missing in the endogenous proviruses (27). It is at this approximate position (2.25 kb from the proviral 5' end) in baboon DNA that we observe the Bam HI heterogeneity (Fig. 3).

Perhaps the most striking feature of the BaEV provirus organization in BEF-3 DNA is that there is a Bam HI recognition site very close to the 5' end of many proviral genomes. Whether this site resides within the genome itself or in the closely adjacent baboon flanking DNA is not known. Since there is no Bam HI site near the 3' end of the integrated genomes (i.e., a 1.4 kb fragment is not released in a Bam HI digest of BEF-3 DNA) as would be expected if the 5' Bam HI site is within the terminal repeat, one would predict that if the 5' Bam HI site is present within the terminal repetition it is the result of a mutation in the 5' repeat but not in the 3' repeat (or vice versa). The observation of a faint 1.25 kb band in Bam HI digests of RD(BAB8-K) DNA (Fig. 1, Table 1) could be explained by the following hypothesis. If the BaEV proviral 5' end contains five of the six nucleotides of the Bam HI recognition sequence, approximately one of every four integration events should generate such a Bam HI site. This possibility is currently being tested.

In summary, we find distinct nucleotide differences in the multiple inte-

grated BaEV genomes of a baboon embryo fibroblast line, BEF-3, and six tissues from four juvenile baboons. The heterogeneity includes a Bam HI and a Pvu II recognition sequence and an approximately 150 nucleotide pair deletion. Accumulation of a subset of provirus genomes of the uninserted (or deleted) variety may correlate with the ability of these genomes to become expressed as infectious particles. We are currently pursuing this possibility by characterizing individual BaEV genomes in baboon cell DNA by recombinant DNA methods and by studying the infectivity and arrangement of proviruses present in the DNA of baboon tumors.

ACKNOWLEDGMENT

This work was supported by Contract NO1-CO-75380 with the National Cancer Institute, Bethesda, Maryland 20205. We wish to thank Dr. Howard Young for a critical reading of this manuscript.

REFERENCES

1. Lueders, K., Kuff, E. (1977). Cell 12, 963-972.
2. Kuff, E., Lueders, K., Scolnick, E. (1978). J. Virol 28, 66-74.
3. Todaro, G., Callahan, R., Rapp, U., DeLarco, J. (1980). Proc. Royal Society, London. In press.
4. McAllister, R., Nicolson, M., Heberling, R., Charman, H., Rice, N., Gilden, R. (1978). In Advances in Comparative Leukemia Research 1977, (Bentvelzen, P., Hilgers, J., and Yohn, D., eds.) Amsterdam, New York: Elsevier/North-Holland Biomedical Press, pp. 135-138.
5. Todaro, G., Benveniste, R., Sherwin, S., Sherr, C. (1978). Cell 13, 775-782.
6. Rabin, H., Benton, C., Tainsky, M., Rice, N., Gilden, R. (1979). Science 204, 841-842.
7. Benveniste, R., Todaro, G. (1977). Proc. Natl. Acad. Sci. USA 74, 4557-4561.
8. Gilden, R., Oroszlan, S., Young, H., Rice, N., Gonda, M., Cohen, M., Rein, A. In Frontiers in Immunogenetics. In press.
9. Doolittle, W., Sapienza, C. (1980). Nature 284, 601-603.
10. Orgel, L., Crick, F. (1980). Nature 284, 604-607.
11. Cooper, G., Temin, H. (1976). J. Virol. 17, 422-430.
12. Cohen, M., Nicolson, M., McAllister, R., Shure, M., Davidson, N., Rice, N., Gilden, R. (1980). J. Virol. 34, 28-39.
13. Cohen, M., Davidson, N., McAllister, R., Nicolson, M., Rice, N. (1978). Cold Spring Harbor Tumor Virus Symposium.
14. McAlliser, R.M., Melnyk, J., Finklestein, J.Z., Adams, E.D., Gardner, M.B. (1969). Cancer Res. 24, 520-526.
15. Southern, E. (1975). J. Mol. Biol. 98, 503-517.
16. Sutcliffe, J.G. (1978). Nucleic Acids Res. 5, 2721-2728.
17. Steffen, D., Weinberg, R. (1978). Cell 15, 1003-1010.
18. Astrin, S. (1978) Proc. Natl. Acad. Sci. USA 75, 5941-5945.
19. Benveniste, R., Liber, M., Livingston, D., Sherr, C., Todaro, G. (1974). Nature 248, 17-20.
20. Todaro, G., Sherr, C., Benveniste, R., Lieber, M., Melnick, J. (1974). Cell 2, 55-61.

21. Copeland, T., Henderson, L., Vanlaningham-Miller, E., Stephenson, J., Smythers, G., Oroszlan, S. Virol. Submitted.
22. Mann, M., Smith, H. (1977). Nucl. Acids, Res. 4, 4211-4221.
23. Pirotta, V. (1976). Nucl. Acids Res. 3, 1747-1760.
24. Hattman, S., Keister, T., Gottehrer, A. (1978). J. Mol. Biol. 124, 701-711.
25. Roy, P., Weissbach, A. (1975). Nucl. Acids Res. 10, 1669-1684.
26. Lavelle, G., Foote, L., Hererling, R., Kalter, S. (1979). J. Virol. 30, 390-393.
27. Cohen, J.C., Shank, P., Morris, V., Cardiff, R., H. Varmus. (1979). Cell 16, 333-345.