JOURNAL OF VIROLOGY, Sept. 1991, p. 5092–5095
0022-538X/91/095092-04$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 65, No. 9

Prevalence of the Archetypal Regulatory Region and Sequence Polymorphisms in Nonpassaged BK Virus Variants

MASSIMO NEGRINI,[1,2]* SILVIA SABBIONI,[2] RAY R. ARTHUR,[3] ANTONELLA CASTAGNOLI,[1] AND GIUSEPPE BARBANTI-BRODANO[2,4]

Interdepartment Center for Cancer Research,[1] Institute of Microbiology,[2] and Biotechnology Center,[4] University of Ferrara, I-44100 Ferrara, Italy, and Department of Immunology and Infectious Diseases, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Maryland 21205[3]

Received 1 April 1991/Accepted 11 June 1991

Since the first isolation and characterization of BK virus (BKV), a number of BKV variants which differ in genomic structure or antigenic determinants have been described. The regulatory region, in particular, the enhancer elements, show the most divergent sequences among different isolates. The structural organization of a putative ancestral prototype or archetype, from which all of the variants are probably derived, has been proposed. By sequencing the regulatory regions of 13 different isolates from the urine of bone marrow transplant recipients, we determined the structures and sequences of BKV variants diffused in the human population. The enhancer region was amplified by polymerase chain reaction to avoid passage in culture, and the product was directly sequenced. The structure most frequently observed is in agreement with the postulated archetype, containing a single enhancer element with no repeats. By sequence analysis we identified four hot spots of nucleotide variation. These variations are consistent with the existence of two consensus sequences. One sequence motif, observed in about 85% of the isolates, is referred to as the archetypal BKV, while a second motif, observed in the remaining 15% of the variants, is highly reminiscent of the AS strain.

BK virus (BKV) is a human papovavirus originally isolated from the urine of an immunosuppressed patient (7). Since then, several BKV variants have been isolated. The genomes of the Dunlop (DUN), MM, and AS strains have been cloned and completely sequenced (22, 26, 29), while the origin of replication and transcriptional control elements have been sequenced in other variants (16, 17, 19, 23, 25, 27). These isolates differ in several ways, but the structural organization of the early-region transcriptional enhancers is the most variable. In the first BKV isolate, the Gardner strain, the enhancer region is composed of three tandem repeats. Enhancers were first described in simian virus 40 (3), where they consisted of two tandem repeats, suggesting that enhancers are typically made up of a number of tandemly repeated elements. However, BKV strains DIK (27), WW (19), and AS (26) have a linear arrangement of the regulatory region, with no repeats, similar to the archetype BKV proposed by Yoshiike and Takemoto (31). Isolation of different BKV variants may answer questions about the origin of the variation and lead to definition of the ancestral prototype or archetype. It was speculated that BKV strains with tandemly repeated enhancers have acquired this structure during passage in culture, with specific arrangements being selected by plaque purification. Indeed, rearrangement in the transcriptional control region of strain WW occurs during passage in cell culture (20). To determine the most common BKV strain in the human population and clarify its archetypal structure, we analyzed the regulatory region of BKV isolates from 13 bone marrow transplant recipients during the posttransplant period. Of the 13 patients, 2 (no. 7 and 12) had hemorrhagic cystitis (2). We avoided passage in culture by using the polymerase chain reaction (PCR) to amplify the region directly, followed by sequencing of the amplified products.

The BKV regulatory region was amplified by using the primers 5′-TTGAGAGAAAGGGTGGAGGC-3′ (BRP-1) and 5′-GCCAAGATTCCTAGGCTCGC-3′ (BRP-2) (Fig. 1), starting from urinary sediments. The urinary sediments were prepared as described previously (1). Pellets were suspended in PCR buffer and boiled for 10 min, and 5-μl samples were added to 45-μl reaction mixture volumes. The primer concentrations, as optimized by checkerboard titration, were 0.5 and 0.2 μM for BRP-1 and BRP-2, respectively. The cycling parameters were 60 s at 94°C, 30 s at 42°C, and 60 s at 72°C for 50 cycles. To sequence the PCR products directly, we produced a single-strand template by carrying out an asymmetric PCR (11), starting with 0.15 ng of the PCR product, 0.3 pmol of the limiting primer, and 10 pmol of the other primer in a 100-μl reaction. The parameters 95°C-60 s, 60°C-30 s, and 72°C-60 s were repeated for 40 cycles. Products of the asymmetric PCR were purified by using the Centricon 30 spin dialysis procedure (21). Sequencing on the single-strand template product of the asymmetric PCR was performed as previously described (11). Of the 13 samples, 12 have a linear arrangement composed of 68-, 39-, and 63-bp elements, in perfect agreement with the proposed putative archetype (31). Only one sample, no. 14, has a 23-bp repeat (from nucleotides 42 to 65 of Fig. 2), within the 68-bp element (Fig. 1). From this structural analysis, we conclude that the organization of the regulatory region most frequently observed in BKV variants infecting human populations is the proposed archetype and that all of the other strains originate by duplications and/or deletions of the prototype sequence. Analysis of short sequence variation among the numerous isolates indicates that four regions are polymorphic (Table 1). Region 1 can be indifferently AA, TT, or TA, while differences in the other three regions are compatible with the existence of two sequence motifs. The most common motif (85% of the isolates) has the following consensus sequences: region 2, 5′-AAGGAAA-3′; region 3, 5′-AAAA-3′; and region 4, 5′-CTGG-3′. The second motif

* Corresponding author.

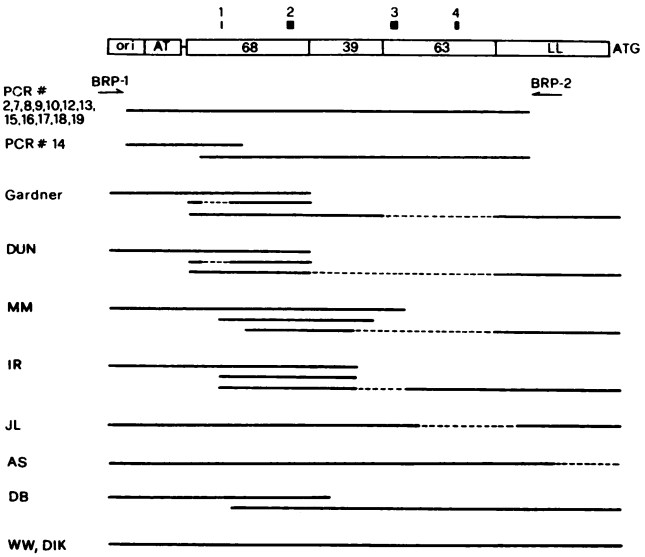

FIG. 1. Schematic representation of the structural organization of the regulatory regions of various BKV isolates. The regions analyzed in this study were enclosed between the two PCR primers BRP-1 (from positions −26 to −7 of reference 28) and BRP-2 (from positions 238 to 219 of Fig. 2) (see the text for sequence details). BKV strains are represented and compared with the archetypal structure proposed by Yoshiike and Takemoto (31; boxed diagram on the top). The region comprises the origin of replication (ori), the A+T-rich sequence (AT), the enhancer repeats composed of the 68-, 39-, and 63-bp elements, and the 63-bp late leader (LL) sequence. The black rectangles on the top of the archetypal sequence indicate the four hot spots for nucleotide variation. Below the diagram of the archetypal region, the structural characteristics of different BKV isolates are indicated: continuous lines denote the presence of the corresponding sequence in the archetypal structure, interrupted lines indicate a deletion, and overlapping segments indicate a duplication. The prevalence of the archetypal structure and how all of the other strains derive from the archetype by sequence duplications and deletions are shown.

(15% of the isolates), similar to the AS strain, has the following consensus sequences: region 2, 5′-AAGCAA-3′; region 3, 5′-AACA-3′; and region 4, 5′-CTGT-3′. Region 2 sequences, 5′-AAGGAAA-3′ and 5′-AAGCAA-3′, are homologous to the adenovirus type 5 E1A enhancer core (12). The strong homology to a regulatory element of another virus suggests that they also have an important function for BKV and that even small differences may determine different biological properties. By comparison of all of the sequences, a consensus sequence (Fig. 2) of the archetypal

```
          10         20         30         40         50         60
          :          :          :          :          :          :
  1 GCCTCGGCCT CTTATATATT ATAAAAAAAA AGGCCACAGG GAGGAGCTGC TTACCCATGG
 61 AATGCAGCCA AACCATGACC TCAGGAAGGA AAGTGCATGA CTGGGCAGCC AGCCAGTGGC
121 AGTTAATAGT GAAACCCCGC CCCTAAAATT CTCAAATAAA CACAAGAGGA AGTGGAAACT
181 GGCCAAAGGA GTGGAAAGCA GCCAGACAGA CATGTTTTGC GAGCCTAGGA ATCTTGGCCT
241 TGTCCCCAGT TAAACTGGAC AAAGGCCATG
```

FIG. 2. Nucleotide sequence of the archetypal BKV regulatory region, deduced as a consensus sequence from comparison of the sequences of all of the BKV strains presented in Fig. 1. Numbering begins at the origin of replication, as in reference 28. Duplication of PCR no. 14 is from nucleotides 42 to 65.

TABLE 1. Hot spots for nucleotide variation and sequence motifs in the regulatory regions of different BKV strains

Strain[a]	Sequence of the designated nucleotides[b] of region[c]:			
	1, 52–53	2, 86–92	3, 145–148	4, 179–182
Gardner	TA	AAGGAAA	–[d]	–
MM	TA	AAGGAAA	AAAA	–
IR	TA	AAGGAAA	–	CTGG
JL	TA	AAGGAAA	GAAA	–
DIK	AA	AAGGAAA	AAAA	CTGG
WW	AA	AAGGAAA	AAAA	CTGG
AS	TA	AAGC–AA	AACA	CTGT
PCR no. 2	TT	ANGNAAA	GAAA	CTGG
7	TA	AAGC–AA	AACA	CTGT
8	TA	AAGGAAA	AAAA	CTGG
9	TA	AAGGAAA	GAAA	CTGG
10	TA	AAGGAAA	NAAA	CTGG
12	AA	AAGGAAA	AAAA	CTGG
13	TT	AAGGAAA	AGAA	GTAG
14	TT	AAGC–AA	AACA	CTTT
15	AA	AAGGAAA	AAAA	CTGG
16	AA	AAGGAAA	AAAA	CTGG
17	TA	AAGGAAA	GAAA	CTGG
18	TA	AAGGAAA	AAAA	CTGG
19	AA	AAGGAAA	AAAA	CTGG
Motifs				
1	TA	AAGGAAA	AAAA	CTGG
2	TA	AAGC–AA	AACA	CTGT

[a] The sequences of strains Gardner, MM, IR, JL, DIK, WW, and AS are from the literature; the sequence of PCR strains is from this study.
[b] Numbering begins at the origin of replication and proceeds towards the late region, as in Fig. 2.
[c] The four regions are schematically represented in Fig. 1.
[d] —, not found in the corresponding strain.

BKV regulatory region can be deduced and used to verify all new BKV isolates.

No heterogeneity was observed in PCR-amplified sequences, suggesting that urine samples contained only viruses with the archetypal organization of the regulatory region. However, the presence of undetectable amounts of rearranged variants cannot be ruled out. As noted by Sundsfjord et al. (24), rearranged variants could initiate replication in cell culture because of the advantage conferred by enhancer repeats in this in vitro environment, explaining the constant isolation of strains with rearranged control regions.

Sequence analysis of the transcriptional control region from noncultured isolates enables determination of the wild-type organization of the region. On the basis of the foregoing data, we propose that the regulatory region of the putative archetype (31) is actually the predominant structure present in the BKV strains that infect the general population. Analogous results have been recently obtained by Flaegstad et al. (6). Deviations from this ancestral structure may reflect different biological features. Biological properties of BKV variants with sequence variations in the regulatory region should be compared with the WW strain, whose sequence is the closest to that of the archetypal BKV deduced from the 13 BKV isolates described. BKV strains with tandemly repeated sequences in the regulatory region most probably arose during passage in cell culture, with specific arrangements being selected by plaque purification. In agreement with this hypothesis, BKV strain WW (19) and other isolates (23) with single-repeat arrays cannot grow in cell culture under conditions that allow growth of other BKV strains. Indeed, forced growth of BKV strain WW in culture deter-

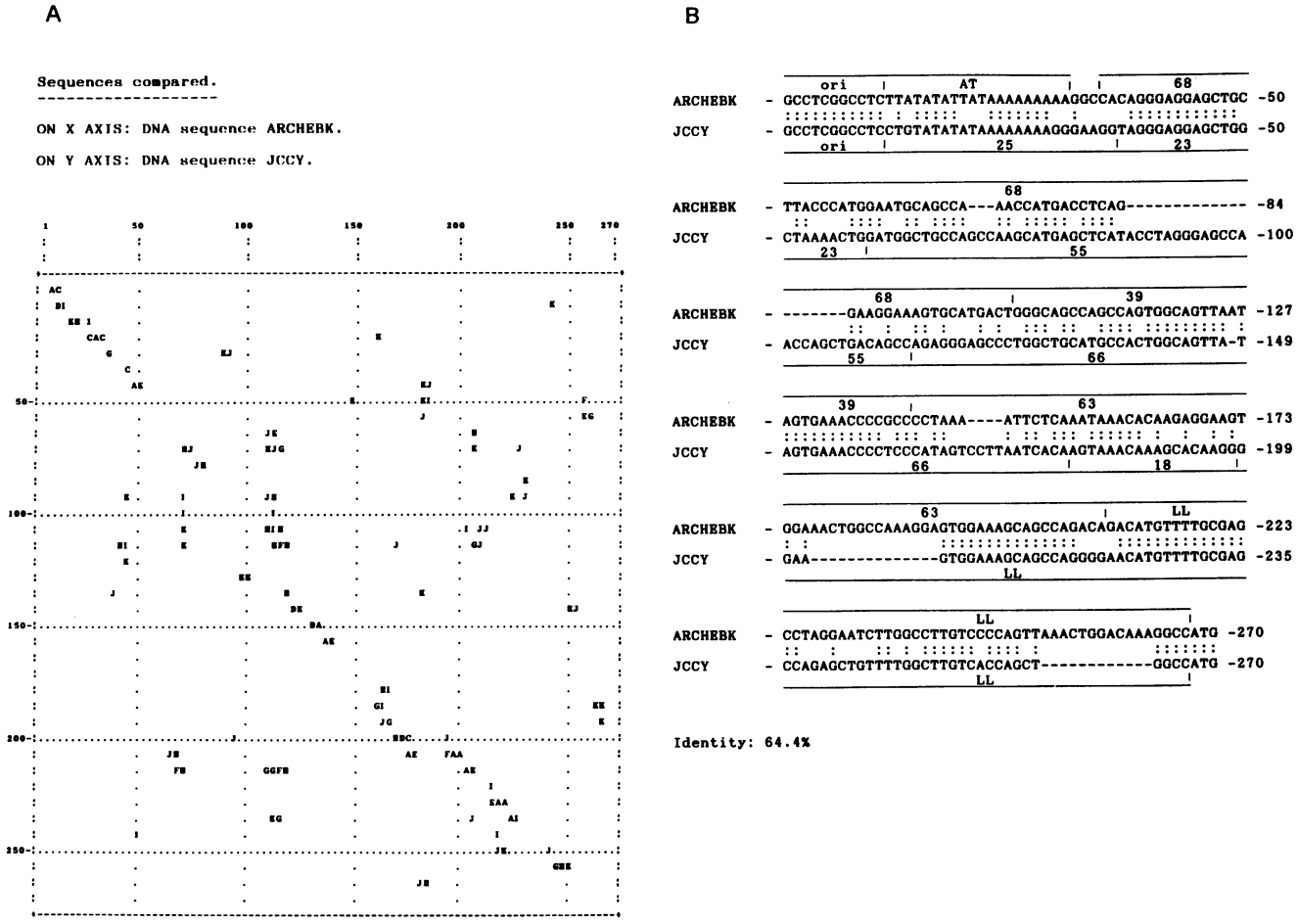

FIG. 3. Comparison of archetypal sequences of human papovaviruses BKV and JCV (30). (A) Matrix analysis was done with the NMATPUS program (which uses Pustell's algorithm [18]) of the PCGENE package, release 6.01 (Intelligenetics, Inc./Genofit S.A.). The parameters used were the following: window size, ±5 bases; minimum match percentage shown, 80%. Symbols: A, 100%; B, 98 to 99%; C, 96 to 97%; D, 94 to 95%; E, 92 to 93%; F, 90 to 91%; G, 88 to 89%; H, 86 to 87%; I, 84 to 85%; J, 82 to 83%; K, 80 to 81%. The interrupted diagonal in the output denotes divergence between two evolutionarily related sequences. (B) A more consistent indication of the conserved homologous regions in the two viruses was obtained by aligning the two sequences. The result was obtained by using the NALIGN program with the algorithm originally developed by Gotoh (8). The origin of replication and the A+T-rich region are highly conserved, as is the BKV 39-bp element in the enhancer region, while the 68-bp element is almost completely divergent. Two other regions (191 to 205 and 210 to 225 bp), at the boundary between the 63-bp element and LL, are 100% conserved between the two viruses.

mines the evolution and propagation of BKV variants with rearranged regulatory regions (20).

Even if the sequences of BKV strains with tandem repeats were all present in the archetypal structure, junctions between repeats or deletions might determine new sequence determinants. In fact, a new binding site for the AP-1 transcription factor was generated by a sequence junction in strain DUN (15). *trans*- and *cis*-acting elements are involved in the regulation of transcription initiation: host cell specificity for both viral and cellular enhancers (9, 13) is controlled by the presence or absence of either positively or negatively acting factors (10, 14). Thus, a different structure of the enhancer region may confer different biological properties on BKV strains harboring variations of the basic archetypal motif, like a new cell specificity. The variant BKV IR was isolated from a human insulinoma, and the genome of the reactivated virus has the same structural features as the viral DNA detected in the tumor (4, 16), suggesting that its structural organization evolved during in vivo persistence and was not due to cell culture propagation. It is possible that the enhancer organization of BKV IR contributes new properties to the virus and is causally related to tumor development. BKV variants similar to IR have been identified and isolated from other human tumors and tumor cell lines (5, 15). The presence of a BKV strain (TU) with a rearranged control region in a human host was also recently reported (24). However, this strain was found only in the urine of Norwegian patients, indicating a probable specific geographic localization for this strain (6).

Recently, a putative archetype of the other human papovavirus JC virus (JCV; CY strain) was identified (30). In this case too, the enhancer element has a linear arrangement with no repeats, indicating that the structure of the region in different strains has followed the same evolutionary pathway as in BKV. Matrix and sequence comparisons between BKV and JCV archetypes (Fig. 3) show a clear relationship. The most divergent sequence is between the 68-bp element of BKV and the 23- and 55-bp elements of JCV (30), which

most likely confer different cell specificities on the two viruses. The region with the highest grade of homology is within the 39-bp element in BKV and the 66-bp element in JCV. It is notable that this region is completely or partially deleted in some BKV variants (Gardner, DUN, and IR) and in JCV strains isolated from patients affected by progressive multifocal leukoencephalopathy (30). Since this region partially overlaps the negative regulatory domain in the BKV enhancer (10), it is tempting to speculate that deletion of this region can determine the evolution of more aggressive strains.

Comparative analysis of the structures of the regulatory regions in different BKV isolates demonstrates no evolutionary relationship within the group, with the exception of the DUN strain, which clearly derives from the Gardner strain. However, from short-sequence analysis, the existence of at least two closely related types of BKV variants can be proposed: one type, corresponding to 85% of the isolates, can be considered the archetypal BKV, and the other 15% of the isolates belong in a second type which includes the AS strain.

Nucleotide sequence accession number. The GenBank accession number for the nucleotide sequence in Fig. 2 is M64605.

We thank Keerti V. Shah for helpful discussions. We are grateful to A. Bevilacqua for preparation of the manuscript.

M.N. and A.C. are recipients of fellowships from the Consorzio Interuniversitario Nazionale per la Bio-Oncologia. S.S. is a recipient of a fellowship from the Associazione Italiana per la Ricerca sul Cancro. This work was supported by the Consiglio Nazionale delle Ricerche, Progetto Finalizzato "Oncologia," and by the Associazione Italiana per la Ricerca sul Cancro.

REFERENCES

1. **Arthur, R. R., S. Dagostin, and K. V. Shah.** 1989. Detection of BK virus and JC virus in urine and brain tissue by the polymerase chain reaction. J. Clin. Microbiol. **27**:1174–1179.
2. **Arthur, R. R., K. V. Shah, P. Charache, and R. Saral.** 1988. BK and JC virus infections in recipients of bone marrow transplants. J. Infect. Dis. **158**:563–569.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299–308.
4. **Caputo, A., A. Corallini, M. P. Grossi, L. Carrà, P. G. Balboni, M. Negrini, G. Milanesi, G. Federspil, and G. Barbanti-Brodano.** 1983. Episomal DNA of a BK virus variant in a human insulinoma. J. Med. Virol. **12**:37–49.
5. **Corallini, A., M. Pagnani, P. Viadana, E. Silini, M. Mottes, G. Milanesi, G. Gerna, R. Vettor, G. Trapella, V. Silvani, G. Gaist, and G. Barbanti-Brodano.** 1987. Association of BK virus with human brain tumors and tumors of pancreatic islets. Int. J. Cancer **39**:60–67.
6. **Flaegstad, T., A. Sundsfjord, R. R. Arthur, M. Pedersen, T. Traavik, and S. Subramani.** 1991. Amplification and sequencing of the control regions of BK and JC virus from human urine by polymerase chain reaction. Virology **180**:553–560.
7. **Gardner, S. D., A. M. Field, D. V. Coleman, and B. Hulme.** 1971. New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet **i**:1253–1257.
8. **Gotoh, O.** 1982. An improved algorithm for matching biological sequences. J. Mol. Biol. **162**:705–708.
9. **Grinnell, B. W., D. T. Berg, and J. Walls.** 1986. Activation of the adenovirus and BK virus late promoters: effects of the BK virus enhancer and *trans*-acting viral early proteins. Mol. Cell. Biol. **6**:3596–3605.
10. **Grinnell, B. W., D. T. Berg, and J. D. Walls.** 1988. Negative regulation of the human polyomavirus BK enhancer involves cell-specific interaction with a nuclear repressor. Mol. Cell. Biol. **8**:3448–3457.
11. **Gyllenstein, U. B., and H. A. Erlich.** 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. Proc. Natl. Acad. Sci. USA **85**:7652–7656.
12. **Hearing, P., and T. Shenk.** 1983. The adenovirus type 5 E1A transcriptional control region contains duplicated enhancer elements. Cell **33**:695–703.
13. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue specific gene expression. Science **236**: 1237–1245.
14. **Markowitz, R.-B., S. Tolbert, and W. S. Dynan.** 1990. Promoter evolution in BK virus: functional elements are created at sequence junctions. J. Virol. **64**:2411–2415.
15. **Negrini, M., P. Rimessi, C. Mantovani, S. Sabbioni, A. Corallini, M. A. Gerosa, and G. Barbanti-Brodano.** 1990. Characterization of BK virus variants rescued from human tumours and tumour cell lines. J. Gen. Virol. **71**:2731–2736.
16. **Pagnani, M., M. Negrini, P. Reschiglian, A. Corallini, P. G. Balboni, S. Scherneck, G. Macino, G. Milanesi, and G. Barbanti-Brodano.** 1986. Molecular and biological properties of BK virus-IR, a BK virus variant isolated from a human tumor. J. Virol. **59**:500–505.
17. **Pater, A., M. M. Pater, L.-S. Chang, K. Slawin, and G. Di Mayorca.** 1983. Multiple origins of the complementary defective genomes of RF and origin proximal sequences of GS, two human papovavirus isolates. Virology **131**:426–436.
18. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination. Nucleic Acids Res. **12**:643–656.
19. **Rubinstein, R., N. Pare, and E. H. Harley.** 1987. Structure and function of the transcriptional control region of nonpassaged BK virus. J. Virol. **61**:1747–1750.
20. **Rubinstein, R., B. C. A. Schoonakker, and E. H. Harley.** 1991. Recurring theme of changes in the transcriptional control region of BK virus during adaptation to cell culture. J. Virol. **65**:1600–1604.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Seif, I., G. Khoury, and R. Dhar.** 1979. The genome of human papovavirus BKV. Cell **18**:963–977. (Erratum, **19**:567, 1980.)
23. **Sugimoto, C., K. Hara, F. Taguchi, and Y. Yogo.** 1989. Growth efficiency of naturally occurring BK virus variants in vivo and in vitro. J. Virol. **63**:3195–3199.
24. **Sundsfjord, A., T. Johansen, T. Flaegstad, U. Moens, P. Villand, S. Subramani, and T. Traavik.** 1990. At least two types of control regions can be found among naturally occurring BK virus strains. J. Virol. **64**:3864–3871.
25. **Tavis, J. E., R. J. Frisque, D. L. Walker, and F. A. White III.** 1990. Antigenic and transforming properties of the DB strain of the human polyomavirus BK virus. Virology **178**:568–572.
26. **Tavis, J. E., D. L. Walker, S. D. Gardner, and R. J. Frisque.** 1989. Nucleotide sequence of the human polyomavirus AS virus, an antigenic variant of BK virus. J. Virol. **63**:901–911.
27. **ter Schegget, J., C. J. A. Sol, E. W. Baan, J. van der Noordaa, and H. van Ormondt.** 1985. Naturally occurring BK virus variants (JL and Dik) with deletions in the putative early enhancer-promoter sequences. J. Virol. **53**:302–305.
28. **Watanabe, S., E. Soesa, S. Uchida, and K. Yoshiike.** 1984. DNA rearrangement affecting expression of the BK virus transforming gene. J. Virol. **51**:1–6.
29. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. Science **206**:456–462.
30. **Yogo, Y., T. Kitamura, C. Sugimoto, T. Ueki, Y. Aso, K. Hara, and F. Taguchi.** 1990. Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J. Virol. **64**:3139–3143.
31. **Yoshiike, K., and K. K. Takemoto.** 1986. Studies with BK virus and monkey lymphotropic papovavirus, p. 295–326. *In* N. P. Salzman (ed.), The Papovaviridae, vol. 1. The polyomaviruses. Plenum Press, New York.