JOURNAL OF VIROLOGY, Feb. 1989, p. 901–911 Vol. 63, No. 2
0022-538X/89/020901-11$02.00/0

Nucleotide Sequence of the Human Polyomavirus AS Virus, an Antigenic Variant of BK Virus

JOHN E. TAVIS,[1] DUARD L. WALKER,[2] SYLVIA D. GARDNER,[3] AND RICHARD J. FRISQUE[1]*

Department of Molecular and Cell Biology, The Pennsylvania State University, University Park, Pennsylvania 16802[1]; Department of Medical Microbiology, University of Wisconsin Medical School, Madison, Wisconsin 53706[2]; and Central Public Health Laboratory, Virus Reference Laboratory, London NW9 5HT, England[3]

Received 15 August 1988/Accepted 2 November 1988

The complete DNA sequence of the human polyomavirus AS virus (ASV) is presented. Although ASV can be differentiated antigenically from the other human polyomaviruses (BK and JC viruses), it shares 94.9% homology at the nucleotide level with the Dunlop strain of BK virus. Differences found in ASV relative to BK virus include the absence of tandem repeats in its regulatory region, the deletion of 32 nucleotides in the late mRNA leader region (altering the initiation codon for the agnoprotein), the presence of a cluster of base pair substitutions within the coding region of the major capsid protein, VP1, and the absence of 4 amino acids in the carboxy-terminal region of the early protein, T antigen. The 43 nucleotides deleted in the Dunlop strain of BK virus relative to the Gardner prototype strain of BK virus are present in ASV. Possible reasons for the distinct antigenicity of the ASV capsid, given the high degree of nucleotide homology with BK virus, are discussed. To reflect the high degree of sequence homology between ASV and BK virus, we suggest ASV be renamed BKV(AS).

The first polyomaviruses discovered in humans, BK virus (BKV) and JC virus (JCV), were isolated in 1971 from urine (13) and brain tissue (36), respectively. Subsequently, a number of BKV and JCV variants have been recovered from their human hosts and characterized. In 1980, Coleman and co-workers (7) identified a polyomavirus in the urine of a pregnant woman (initials AS) that replicated in human fetal kidney, lung, and glial cells. Although it exhibited the same host range that BKV did, this new virus did not react with antisera raised against intact BKV virions and was therefore designated AS virus (ASV). In 1983, Gibson and Gardner (14) examined the restriction enzyme patterns and antigenic characteristics of ASV. Unique *Bam*HI and *Hin*dIII restriction profiles were observed, but overall the restriction patterns of ASV DNA were similar to those of BKV DNA. When the antigenic properties of the ASV capsid were analyzed with type-specific rabbit antisera raised against BKV or JCV virions, only a very weak cross-reaction was detected between BKV and ASV; no cross-reaction was observed between JCV and ASV. A sero-epidemiological study of 350 people demonstrated that 40 to 50% of the adult English population possess antibodies to ASV and that seroconversion generally occurs between the ages of 10 and 20 (14). Initial BKV infections generally occur at a younger age and in a higher percentage (70 to 80%) of the people (12). Gibson and Gardner (14) concluded from this data that ASV was a new human polyomavirus and that it was more closely related to BKV than to JCV.

Because ASV remains an incompletely characterized human virus that is endemic in the population, we have determined its nucleic acid sequence. This information has allowed us to analyze the genetic organization of the viral genome and to study its evolutionary relationships with BKV and JCV. The restriction fragment sizes determined by Gibson and Gardner (14) are confirmed by sequence data, and the cloned ASV possesses the same serological reactivity and lytic activity as the original isolate, indicating that the sequenced ASV strain is the same as the one originally studied by these investigators.

The sequence of ASV reveals a remarkable similarity to that of the Dunlop strain of BKV [BKV(DUN); 44]; the two viruses are 94.9% homologous at the nucleotide level. Although ASV and BKV(DUN) differ considerably in their noncoding regulatory regions, the organization of the ASV regulatory signals is similar to that found in some BKV variants. Differences between the ASV and BKV(DUN) DNAs are also apparent within those sequences encoding the carboxy terminus of T antigen and the amino terminus of the agnoprotein. Overall, the similarity of ASV and BKV(DUN) at the sequence level is in contrast to the variance previously reported for their restriction patterns and antigenic properties. The sequence analysis of ASV indicates that it is not a new human polyomavirus but an antigenic variant of BKV.

MATERIALS AND METHODS

Viruses and plasmids. The passage history of ASV has included propagation in primary human fetal glial cells (2 passes), human embryonic kidney cells (3 passes), and human embryonic lung cells (9 passes). The virus was present originally in the primary human fetal glial and human embryonic kidney cells as a mixed infection with JCV (14). Only JCV was expressed in these cells, and it was not until the mixture was passaged in human embryonic lung cells that JCV expression was replaced by that of ASV. Separation from JCV was achieved by a series of terminal dilutions of the viruses in human embryonic lung cells (14). DNA was isolated from this preparation of virions, cut at the unique *Eco*RI site in the late coding region, and cloned into pBR322. To ensure that these DNAs represented viable virus, viral sequences were removed from the plasmid, recircularized, and transfected into WI-38 cells (a human diploid cell line established from human embryonic lung cells) by the modified DEAE-dextran transfection method (47). Transfected cells were inspected for the presence of nuclear viral T antigen by indirect immunofluorescence (5) and for cyto-

* Corresponding author.

pathic effect (34). The production of infectious virions was confirmed by passage of extracts of the transfected cells onto additional WI-38 cells and looking for the appearance of viral T antigen, cytopathic effect, and ASV DNA (determined by restriction enzyme digests of low-molecular-weight DNA) (19). One clone, pASV2, was chosen for further analysis. Restriction endonuclease patterns, serological reactivity, and lytic activity all indicate that pASV2 is a representative clone of the original ASV population. Viral sequences were excised from pBR322 and ligated into the *Eco*RI site of the plasmid sequencing vector Bluescript-minus, KS polylinker (BSMKS; Stratagene Cloning Systems) to generate the plasmid pASV2-BSMKS.

Unidirectional deletions. To produce subclones of pASV2-BSMKS for sequencing, unidirectional deletions were generated by the method of Henikoff (18) as adapted by the Erase-a-Base system (Promega Biotec) and further modified in our hands. Briefly, 15 μg of pASV2-BSMKS was cut to completion with *Apa*I and *Cla*I (New England BioLabs, Inc.) and treated with 500 U of exonuclease III (Exo III; Stratagene Cloning Systems). Aliquots containing 0.6 μg of pASV2-BSMKS were removed from the Exo III digestion at 30-s intervals and placed on dry ice. Because *Apa*I generates a 3′ overhang that is resistant to Exo III digestion, adjacent plasmid sequences (including the sequencing primer sites) were protected and digestion proceeded only from the susceptible 5′ overhang generated by *Cla*I. Progressively longer single-stranded regions were generated with increasing Exo III digestion times. These were removed by digestion with 1.6 U of S1 nuclease (Pharmacia Fine Chemicals) per time point. The partially deleted plasmids were purified, and their ends were filled in by using Klenow fragment and deoxynucleoside triphosphates. Plasmids were recircularized and used to transform *Escherichia coli* NM522 cells. Plasmid DNA was isolated from bacterial colonies by rapid alkaline extraction (2) and screened for deletions of appropriate sizes by restriction enzyme digests.

Template preparation and sequencing. Double-stranded plasmid templates for sequencing were generated by a modification of the rapid alkaline extraction procedure (2). The double-stranded template sequencing protocol was based on that of Chen and Seeburg (4) as modified by New England BioLabs; [α-^{32}P]dATP was used to radiolabel the extension products. The entire sequence was determined at least twice, and 95% of the sequence was determined for both strands. Sections not determined on both strands by the enzymatic method were confirmed by the chemical degradation method of Maxam and Gilbert (31).

DNA sequence analysis. Sequence data was compiled by using an IBM Personal Computer XT coupled with the International Biotechnologies, Inc., Gel Reader. Analysis software was the IBI DNA/Protein Sequence Analysis System by J. Pustell. Hydropathy indexes were generated by the software by using the scale of Kyte and Doolittle (26). The hydropathy index is a measure of the average hydrophobicity in the center of a sliding window 9 amino acids (aa) in length. It is a dimensionless value ranging from −4 (very hydrophilic) to 4 (very hydrophobic), with neutral hydrophobicity at −0.4.

Transformation of Rat 2 cells. Transforming activity of ASV was determined by the dense-focus assay by using a subclone of Rat 2 cells (2a, 50). Cells (4×10^5) were seeded onto 60-mm tissue culture dishes and propagated in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum. Cells were transfected 12 to 16 h later with 1.0 μg of plasmid DNA by the modified calcium phosphate method (52). After transfection, cultures were maintained in Dulbecco modified Eagle medium with 5% fetal bovine serum and medium was changed every 4 to 5 days. Cells were fixed with 3.7% formaldehyde and stained with hematoxylin at 30 to 38 days posttransfection to facilitate enumeration of dense foci. Cells from three independent foci were isolated and cloned by limited dilution for each DNA used in experiment 2 (see Table 2). These cells exhibited a transformed morphology and expressed nuclear T antigen.

Immunological methods. Hemagglutination (HA) and hemagglutination inhibition (HAI) titers were determined as described previously (35), except that a microdilution technique was employed (46). The highest dilution of virus suspension that showed complete hemagglutination was considered to contain 1 HA unit (HAU) of virus. Antibodies were produced by injecting adult female New Zealand white rabbits intravenously with ASV, BKV, or JCV. Animals received a single injection and were bled 10 days later. Human sera were obtained from R. Golubjatnikov, Immunology Section, Wisconsin State Laboratory of Hygiene. All sera were heated at 56°C for 30 min and treated with $NaIO_4$ before being tested for HAI antibodies (35). An HAI titer of 32 or greater was considered a positive indicator of prior infection by the virus.

RESULTS

ASV genome. The sequence of 5,098 base pairs (bp) representing the ASV genome is shown in Fig. 1. The numbering system begins at the center of the DNA origin of replication and proceeds towards the late region of the genome, following the convention of Fiers et al. (10) for simian virus 40 (SV40), Seif et al. (45) for BKV(DUN), and Frisque et al. (11) for JCV.

The nucleotide sequence of ASV reveals an unexpectedly high degree of homology (94.9%) with that of BKV(DUN) (44), given that previous immunological and restriction enzyme analyses by Gibson and Gardner (14) had indicated only a weak relationship between ASV and the prototype BKV. The genetic organization of ASV is essentially identical to that of SV40, BKV, and JCV; open reading frames for two early (T, t) and four late (VP1, VP2, VP3, and agnoprotein) proteins are predicted from the sequence data.

Regulatory region. The 319 nucleotides spanning positions 4990 to 210 represent the noncoding regulatory region of ASV. Sequence comparisons in the regulatory region are between ASV and the prototype BKV [BKV(WT)] isolated by Gardner (13), because BKV(WT) contains 43 bp not

FIG. 1. Nucleotide sequence of ASV. The circular genome of ASV consist of 5,098 bp. Numbering begins near the center of the presumed origin of replication (within large box) and proceeds towards the late region. The strand listed has the polarity of the late mRNAs. The proposed coding regions for the ASV proteins are shown to the right of the sequence. Initiation and termination codons are indicated by small boxes. DT, 5′ mRNA splice donor site for the T-antigen mRNA; Dt, donor site for the t-antigen mRNA; AE, shared early 3′ mRNA acceptor site; DL, shared 5′ mRNA splice donor site for the late messages; A1, 3′ splice acceptor site for the VP1 mRNA; A2/3, acceptor site for the common VP2 and VP3 mRNA; ——, potential polyadenylation signals; – – – –, AT-rich region; △, location of a 32-bp deletion in ASV relative to BKV. The 67-, 39-, and 63-bp blocks are delineated by brackets and numbers below the sequence.

A

```
         20         40         60         80        100
GCCTCGGCCT CTTATATATT ATAAAAAAAA AGGCCACAGG GAGGAGCTGC TTACCCATGG AATGCAGCCA AACCATGACC TCAAGAAGCA AGTGCATGAC
                                       67
        120        140        160        180        200
TGGGCAGCCA GCCAGTGGCA GTTAATAGTG AAACCCCGCC CCTAACATTC TCAAATAAAC ACAAGAGGAA GTGGAAACTG TCCAAAGGAG TGGAAAGCAG
 39                                        63
        220        240        260        280        300
CCAGACAGAC ATGTTTTGCG AGCCTAAGAA TCTTGTGGTT TTGCGCCAGC TGTCACGACA AGCTTCAGTG AAAGTTGGTA AAACCTGGAC TGGAACTAAA
        320        340        360        380        400
AAAAGAGCTC AGAGGATTTT TATTTTTATT TTAGAGCTTT TGCTGGAATT TTGTAGAGGT GAAGACAGTG TAGACGGGAA AAACAAAAGT ACCACTGCTT
        420        440        460        480        500
TACCTGCTGT AAAAGACTCT GTAAAAGACT CCTAGGTAAG TAATGCTTTT TTTTTGTATT TTCAGGTTGA TGGGTGCTGC TCTAGCACTT TTGGGGGACC
                                    DL                    A2/3
        520        540        560        580        600
TAGTTGCCAG TGTATCTGAG GCTGCTGCTG CCACAGGATT TTCAGTGGCT GAAATTGCTG CTGGGGAGGC TGCTGCTGCC ATAGAAGTTC AAATTGCATC
        620        640        660        680        700
CCTTGCTACT GTAGAGGGCA TAACAACTAC CTCAGAGGCT ATAGCTGCTA TAGGCCTAAC ACCTCAAACA TATGCTGTAA TTGCTGGTGC TCCAGGGGCT
        720        740        760        780        800
ATTGCTGGGT TTGCTGCTTT AATTCAAACT GTTACTGGTA TTAGTTCTTT GGCTCAAGTA GGGTATAGGT TTTTTAGTGA TTGGGATCAC AAAGTTTCCA
        820        840        860        880        900
CTGTAGGCCT TTATCAGCAA TCAGGCATGG CTTTGGAATT GTTTAACCCA GATGAGTACT ATGATATTTT GTTTCCTGGT GTAAATACTT TTGTTAATAA
        920        940        960        980       1000
TATTCAATAT CTAGATCCTA GGCATTGGGG TCCTTCTTTG TTTGCTACTA TTTCCCAGGC TTTGTGGCAT GTTATTAGAG ATGATATACC TGCTATAACT
       1020       1040       1060       1080       1100
TCACAAGAAT TGCAAAGGAG AACAGAGAGA TTTTTTAGGG ACTCTTTGGC TAGATTTTTG GAAGAAACCA CCTGGACAAT TGTAAATGCC CCCATAAACT
       1120       1140       1160       1180       1200
TTTATAATTA TATTCAGGAT TATTATTCTA ATTTGTCCCC TATTAGGCCT TCAATGGTTA GGCAAGTAGC TGAAAGGGAA GGTACCCATG TAAATTTTGG
       1220       1240       1260       1280       1300
CCATACCTAC AGCATAGATA ATGCTGACAG TATAGAAGAA GTTACCCAAA GAATGGATTT AAGAAATAAG GAAAGTGTAC ATTCAGGAGA GTTTATAGAA
       1320       1340       1360       1380       1400
AAAACTATTG CCCCAGGAGG TGCTAATCAA AGAACTGCTC CTCAATGGAT GTTGCCTTTG CTTCTAGGCC TGTACGGGAC TGTAACACCT GCTCTTGAAG
                                                                  A1
       1420       1440       1460       1480       1500
CATATGAAGA TGGCCCCAAC CAAAAGAAAA GGAGAGTGTC CAGGGGCAGC TCCCAAAAAG CCAAAGGAAC CCGTGCAAGT GCCAAAACTA CTAATAAAAG
       1520       1540       1560       1580       1600
GAGGAGTAGA AGTTCTAGAA GTTTAAAACTG GGGTAGATGC TATAACAGAG GTAGAATGCT TTCTAAACCC AGAAATGGGG GATCCAGATG ATAACCTTAG
       1620       1640       1660       1680       1700
GGGCTATAGT CAGCACCTAA GTGCTGAAAA TGCCTTTGAG AGTGATAGCC CAGACAGAAA AATGCTTCCT TGTTACAGTA CAGCAAGAAT TCCACTGCCC
       1720       1740       1760       1780       1800
AACCTAAATG AGGACCTAAC CTGTGGAAAT CTACTAATGT GGGAGGCTGT AACTGTAAAA ACAGAGGTTA TTGGAATAAC TAGCATGCTT AACCTTCATG
       1820       1840       1860       1880       1900
CAGGGTCCCA AAAAGTTCAT GAGAATGGTG GAGGTAAACC TGTCCAAGGC AGTAATTTCC ACTTTTTTGC TGTGGGTGGA GACCCCTTGG AAATGCAGGG
       1920       1940       1960       1980       2000
AGTGCTAATG AATTACAGAA CAAAGTACCC ACAAGGTACT ATAACCCCTA AAAACCCTAC AGCTCAGTCC CAGGTAATGA ATACTGATCA TAAGGCCTAT
       2020       2040       2060       2080       2100
TTGGACAAAA ACAATGCTTA TCCAGTTGAG TGCTGGATTC CTGATCCTAG TAGAAATGAA AATACTAGGT ATTTTGGAAC TTACACAGGA GGGGAAAATG
       2120       2140       2160       2180       2200
TTCCTCCAGT ACTTCATGTT ACCAACACAG CTACCACAGT GTTGCTGGAT GAACAGGGTG TGGGGCCTCT GTGTAAAGCT GATAGCCTGT ATGTTTCAGC
       2220       2240       2260       2280       2300
TGCTGATATT TGTGGGCTGT TTACTAACAG CTCTGGGACA CAACAGTGGA GAGGCCTTGC AAGATATTTT AAGATTCGCC TGAGAAAAAG ATCTGTGAAG
       2320       2340       2360       2380       2400
AATCCTTACC CAATTTCCTT TTTGCTAAGT GACCTTATAA ACAGGAGAAC CCAAAAAGTG GATGGGCAGC CTATGTATGG TATGGAATCT CAGGTTGAGG
       2420       2440       2460       2480       2500
AGGTAAGGGT GTTTGATGGC ACAGAACAGC TTCCAGGGGA CCCAGATATG ATAAGATATA TTGACAGACA AGGACAATTG CAAACAAAAA TGGTTTAAAC
```

Agnoprotein

VP2

VP2/3

VP1

B

2520 2540 2560 2580 2600
AGGTGCTTTA TTGTACATAT ATATGCTTAA TAAATGCTGC TTTTGTATAA CACAGTTGAA GCTTCTGTTA TTTTGGGGGT GGTGTTTTAG GCCTTTTAAA

2620 2640 2660 2680 2700
ACACTGAAAG CCTTTACACA AATGTAACTC TTGGCTGTGA GGGTTTTCTG AATCAGGGGC TGAAGTATCT GAGACTTGGG AAGAGCATTG TGATTGGGAT

2720 2740 2760 2780 2800
TCAGTGCTTG ATCCATGTCC AGAGTCTTCA GTTTCTGAAT CTTCTTCTCT TGTAATATCA AGAATACATT TTCCCATGCA TATATTATAT TTCATCCTTG

2820 2840 2860 2880 2900
AAAAAGTATA CATACTTATC TCAGAATCCA GCCTTTCCTT CCATTCAACA ATTCTAGACT GTATATCTTT TGAAAAATCA GCTACAGGCC TAAACCAAAT

2920 2940 2960 2980 3000
TAGTAGTAGC AAAAGGGTCA TTCCACTTTG TAATATTCTT TTTTCAAGTA AAAACTCAGA GTTTTGCAGG GACTTTCTTA AATATATTTT GGGTCTAAAA

3020 3040 3060 3080 3100
TCTATCTGTC TTACAAATCT AGCCTGAAGA GTTTTAGGGA CAGGATACTC ATTCATTGTA ACTAACCCTG GTGGAAATAT TTGTGTTCTT TTGTTTAAAT

3120 3140 3160 3180 3200
GTTTCTTTTC TAAATTAACC TTAACACTTC CATCTAGATA ATCCCTCAAA CTGTCTAAAT TGTTTATTCC ATGTCCTGAA GGCAAATCCT TTGATTCAGC

3220 3240 3260 3280 3300
TCCTGTCCCT TTTACATCTT CAAAAACAAC CATGTACTGA TCAATAGCCA CACCCAGTTC AAAAGTTAGC CTTTCCATGG GTAAATTTAC ATTTAAAGCT

3320 3340 3360 3380 3400
TTACCTCCAC ATAAGTCTAA TAACCCTGCA GCTAAGGTTG TTTTGCCACT ATCAATTGGA CCTTTAAATA ACCAGTATCT TCTTTTAGGT ACATTAAAAA

3420 3440 3460 3480 3500
CAACACAGTG AAGAAAATCA AAAATAACAG AATCCATTTT AGGTAGCAAA CAATGTAGCC AAGCAACCCC TGCCATATAT TGTTCTAGTA CAGCATTTCC

3520 3540 3560 3580 3600
ATGAGCTCCA AATATTAAAT CCATTTTATC TAATATATGA TTAAATCTTT CTGTTAGCAT TTCTTCCCTG GTCATATGAA GGGTATCTAC TCTTTTTTTA

3620 3640 3660 3680 3700
GCTAATACTG TATCTACTGC TTGCTGACAA ATACTTTTTT GATTTTTACT TTCTGCAAAA ATAATAGCAT TTGCAAAATG CTTTTCATGA TACTTAAAGT

3720 3740 3760 3780 3800
GGTAAGGTTG ATCTTTTTTT TGACACTTTT TACACTCCTC TACATTGTAT TGAAATTCTA AATACATACC CAATAATAAA AACACATCCT CACACTTTGT

3820 3840 3860 3880 3900
TTCTACTGCA TATTCAGTAA TTAATTTCCA AGACACCTGC TTTGTTTCTT CAGGCTCCTC TGGGTTAAAG TCATGCTCCT TTAAGCCCCC TTGAATGCTT

3920 3940 3960 3980 4000
TCCTCTATTA TATGGTATGG ATCCCTAGTT AAGGCACTGT ATAGTAAGTA TTCCTTATTA ACACCCTTAC AAATTAAAAA ACTAAAAGTA CACAGCTTTT

4020 4040 4060 4080 4100
GACAGAAATT ATTAATTGCA GAAACTCTAT GTCTATGTGG AGTTAAAAAG AATATAATAT TATGACCAGC ACACATGTGT CTACTGATAA AAGTTACAGA

4120 4140 4160 4180 4200
ATATTTTTCC ATAAGTTTTT TATACAGAAT TTGAGCTTTT TCTTTAGTGG TATACACAGC AAAACAGGCA AGTGTTCTAT TACTAAATAC AGCTTGACTA

4220 4240 4260 4280 4300
AGAAACTGGT GTAGATCAGA GGGAAAGTCT TTAGGGTCTT CTACCTTTCT TTTTTTTTTG GGTGGTGTTG AGTGTTGGGA ATCTGCTGTT GCTTCTTCAT

4320 4340 4360 4380 4400
CACTGGCAAA CATATCCTCA TGGCAGAATA AATCTTCATC CCATTTTTCA TTAAAGGAGC TCCACCAGGA CTCCCACTCT TCTGTTCCAT AGGTTGGCAC

T Antigen

AE

4420 4440 4460 4480 4500
CTATAAAAAA AAAATAATTA CTTAGGGTCT TCTTTTAATT TACTACTTTT CTAAATATAA ATTAGTTACC TTAAAGCTTT AGATCTCTGA AGGGAGTTTC

Dt

4520 4540 4560 4580 4600
TCCAATTATT TGGACCCACC ATTGCAGGGT TTCTTCAGTG AGGTCTAAGC CAAACCACTG TGTGAAGCAA TCAATGCAGT AGCAATCTAT CCAAACCAAT

4620 4640 4660 4680 4700
GGCTCTTTTC TTAAAAATTT TCTATTTAAA TGCCTTAATC TTAGCTGACA TAGCATGCAA GGGCAATGCA CTGAAGGCTT TTTGGAACAA ATAGGCCATT

4720 4740 4760 4780 4800
CCTTGCAGTA CAAAGTATCT GGGCAAAGAG GAAAATCAGC ACAAACCTCT GAGCTATTCC AGGTTCCAAA ATCAGGCTGA TGAGCTACCT TTACATCCTG

DT

4820 4840 4860 4880 4900
CTCCATTTTT TTATATAAAG TATTCATTCT CTTCATTTTA TCCTCGTCGC CCCCTTTGTC AGGGTGAAAT TCCTTACACT TTTTTAAATA GGCTTTTCTC

4920 4940 4960 4980 5000
ATTAAGGGAA GGTTTCCCCA GGCAGCTCTT TCAAGGCCTA AAAGGTCCAT GAGCTCCATG GATTCTTCCC TGTTTAAGAC TTTATCCATT TTTGCAAAAA

t Antigen

5020 5040 5060 5080
ATTGCAAAAG AATAGGGATT TCCCCAAATA GTTTTGCTAG GCCTCAGAAA AAGCCTCCAC ACCCTTACTA CTTCAGAGAA AGGGTGGAGG CAGAGGCG

FIG. 1—*Continued.*

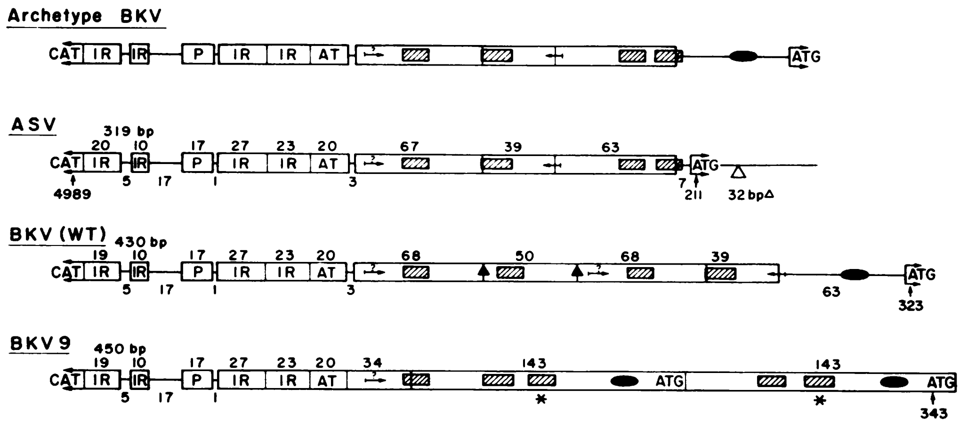

FIG. 2. Comparison of the archetype BKV, ASV, BKV(WT), and BKV9 regulatory regions. The letters CAT within the open box to the left represent the initiation codon (opposite polarity) for the early proteins. To the far right is the ATG initiation codon for the agnoprotein. Comparisons among the viral DNAs include inverted repeats (IR), AT-rich regions (AT), palindromes (P), and putative enhancer blocks. Within these blocks, NF-BK sites (▨), AP1 sites (▲), L1 sites (⬬), and Sp1 sites (←) (left-pointing arrow indicates CCCGCC polarity, right-pointing arrow indicates opposite polarity) are indicated. ?, Sp1 sites determined solely by similarity to the consensus sequence; *, NF-BK sites formed by sequences unique to BKV9 and determined by similarity to the consensus sequence; △, position of a 32-bp deletion relative to the other viruses. Numbers above the linear arrangement refer to the sizes (in base pairs) of the indicated structures; numbers below refer to the distances between structures.

found in BKV(DUN) (55). As the entire BKV(WT) sequence has not been published, the BKV(WT) numbering system used here begins in the center of the origin of replication and proceeds towards the late coding region [as for BKV(DUN)] and assumes that there are no insertions or deletions in BKV(WT) relative to BKV(DUN) outside of the regulatory region. The ASV regulatory region shows a number of distinct alterations from that of BKV(WT), especially to the late side of the origin of replication. These changes are shown schematically in Fig. 2. The changes discussed below use the terms insertion and deletion with respect to BKV(WT), but they are used for convenience only and are not meant to imply any direct evolutionary descent of ASV from BKV(WT). In all probability the two viruses are derived from a common ancestor and not from each other. Specific minor changes found in ASV relative to BKV(WT) were an extra A in the poly(A) stretch from 4996 to 5001 (ASV nucleotide numbers), a G to C transversion at 5074, a G to A transition at 84, a G to C transversion at 89, a deleted A between 90 and 91, and a G to A transition at 227. Major changes included a 118-bp deletion corresponding to the first and second tandem repeats of BKV(WT) (68–50 bp, BKV(WT) 35 to 152), an insertion of 63 bp after BKV(WT) 259 (ASV 141 to 203), and a 32-bp deletion after ASV 235 corresponding to BKV(WT) 292 to 323.

A block homologous to the 63-bp insertion has been observed before in the BKV variants BKV(Dik) (49), BKV(WW) (42), and BKV(R2) (38). The early 34 bp of the 63-bp block are found in BKV(JL) (49), and the late 47 bp of this block are found in BKV(IR) (37). In none of these cases is the 63-bp block identical to the ASV block. ASV 146 is a C, but the corresponding nucleotide is an A in BKV(Dik), BKV(WW), BKV(R2), and BKV(JL) [it is deleted in BKV(IR)]; ASV 181 is a T, but the corresponding nucleotide is a G in BKV(Dik), BKV(WW), BKV(R2), and BKV(IR) [it is deleted in BKV(JL)]; ASV 144 is an A, but the corresponding nucleotide is a G in BKV(JL) and BKV(R2); and BKV(IR) carries an additional A after ASV 189.

The 32-bp deletion in ASV occurs in the late-leader sequences and affects the position of the ATG initiation codon for the agnogene, as predicted on the basis of comparisons with the other polyomavirus agnoproteins. The deletion removes the A of this ATG and shifts a potential in frame initiation codon 24 bp to the 5′ side of the expected site. If this initiation codon is utilized, an additional 8 aa would be added to the amino-terminal end of the ASV agnoprotein relative to the other polyomavirus agnoproteins.

Early coding region. An analysis of reading frames and comparisons with BKV(DUN) sequences indicate that the shared ATG initiation codon for large and small T antigen is located at 4989, the T-antigen splice donor site at 4747, the t-antigen termination codon TAA at 4471, the t-antigen splice donor site at 4470, the shared splice acceptor site at 4400, and the T-antigen TAA codon at 2568. The consensus sequence AAUAAA$(N)_{10-20}$CA (40) for the eucaryotic mRNA polyadenylation signal is found at nucleotide position 2512 to 2491 (AAUAAAN_{14}CA; polarity of the early ASV mRNA). The early coding region is highly homologous with that of BKV(DUN). Specific alterations in the proposed amino acid sequences for all proteins are in Table 1. The proposed sequences for ASV T and t antigens are 98.1 and 98.8% homologous, respectively, to the corresponding BKV(DUN) proteins.

The sequence data predict an ASV T antigen of 691 aa. This is 4 aa fewer than that of the BKV(DUN) T antigen because of a deletion of 12 bp [BKV(DUN) 2694 to 2705] in the extreme carboxy-terminal region of the ASV T-antigen open reading frame. Point mutations and altered codons at the deletion junctions enlarge the region altered in ASV relative to BKV(DUN) to 8 aa (DUN aa 664 to 671; ASV aa 664 to 667). This portion of T antigen appears to be highly variable in the primate polyomaviruses. Relative to the BKV sequences, JCV has a 4-aa deletion (following JCV aa 666) that partially overlaps the region altered in ASV. SV40 has a 1-aa deletion and 5 nonhomologous amino acids in the corresponding region (SV40 aa 675 to 681) (11).

Late coding region. The late coding region of ASV extends from 211 to 2498. The proposed open reading frame for the agnogene was found to extend from 211 to 435. The open reading frames for the minor capsid proteins VP2 and VP3 extend from 470 to 1525 and from 827 to 1525, respectively. The proposed ASV agnoprotein, VP2, and VP3 are 98.5,

TABLE 1. Proposed coding changes in ASV relative to BKV (DUN)

Protein	Position[a]	Amino acid in:	
		BKV (DUN)	ASV
Early			
T antigen	36	Arg	Lys
	78	Ser	Asn
T antigen	36	Arg	Lys
	78	Ser	Asn
	245	Thr	Ile
	260	Ser	Asn
	414	Ile	Val
	591	Ala	Ser
	592	Thr	Lys
	664	Ala	*[b]
	665	Glu	*
	668	Gln	*
	669	Arg	*
	670	Ser	*
	671	Asp	*
Late			
Agnoprotein	1	Met	Val[c]
VP1	61	Glu	Asp
	66	Phe	Tyr
	68	Leu	Gln
	69	Lys	His
	75	Asp	Ala
	77	Ser	Glu
	82	Glu	Asp
	117	Gln	Lys
	139	His	Asn
	145	Ile	Val
	158	Glu	Asp
	171	Ser	Thr
	175	Asp	Gln
	210	Val	Ile
	219	Ala	Thr
	225	Phe	Tyr
	316	Arg	Lys
	340	Arg	Gln
	353	Lys	Arg
	362	Leu	Val
VP2/VP3	53[d]	Ser	Thr
	89	Ser	Thr
	100	Lys	Arg
	103	Asp	Ser
	175	Ser	Ala
	217	Gln	Asp
	240	Arg	His
	242	His	Asn
	251	Asp	Asn
	267	Gln	Lys
	268	Gln	Glu

[a] Amino acid numbers are from the BKV(DUN) sequence (44).
[b] *, Amino acids deleted; positions 668 to 671 replaced with Glu-Asn.
[c] Preceded by Met-Phe-Cys-Glu-Pro-Lys-Asn-Leu in ASV.
[d] VP2 numbers; VP3 amino acid number 1 is VP2 120.

96.9, and 97.4% homologous at the amino acid level, respectively, to their BKV(DUN) counterparts. As with SV40, BKV, and JCV, the VP3 sequence is a subset of the VP2 sequence. The proposed amino acid sequence for VP2 contains 351 aa; the carboxy-terminal 232 aa are shared with VP3. The open reading frame for the major capsid protein VP1 was found to extend from 1410 to 2498. As with the other polyomaviruses, there are two potential initiation codons for VP1 in the same reading frame. We followed the convention used for BKV (44) and JCV (11) and designated the second ATG as the initiation codon.

Low-resolution S1 nuclease mapping of the late messages in BKV(DUN) indicates that BKV produces a 16S and a 19S late mRNA (28). Unlike SV40, BKV uses only one 5′ leader region for its late messages (28). A single late 5′ leader is also proposed for JCV (11). The late splice sites proposed for BKV(DUN) (44) are found in ASV. The shared splice donor site is at ASV 435, the 19S splice acceptor site is at 466, and the 16S splice acceptor site is at 1368. The proposed polyadenylation signal for both late messages in ASV is $AAUAAA(N)_{16}CACA$ and is at position 2529 to 2554.

There are 11 predicted coding changes in VP2 between ASV and BKV(DUN); 7 of these changes are also in VP3. It was recently reported (54) that SV40 VP2 and VP3 carry a nuclear localization signal near their carboxy end (aa 317 to 323 in VP2, sequence Pro-Asn-Lys-Lys-Lys-Arg-Lys). In the analogous region (VP2 aa 316 to 322), ASV and BKV(DUN) are identical and have the sequence Pro-Asn-Gln-Lys-Lys-Arg-Arg.

The proposed ASV VP1 sequence is 94.5% homologous with BKV(DUN) VP1, the lowest homology of any ASV protein to its BKV(DUN) counterpart. The 8 aa deleted in the JCV protein relative to BKV(DUN) (aa 11 to 18) are present in ASV (11). Of the 20 alterations in the proposed VP1 aa sequence between ASV and BKV(DUN), 7 are clustered from aa 61 to 82 in the amino-terminal quarter of VP1. To begin to assess the possible effects of these changes on the antigenicity of VP1, we plotted the hydropathy profiles for the VP1 proteins of ASV and BKV(DUN) by the method of Kyte and Doolittle (26). There is a strong correlation between hydrophilic regions of proteins and external position in the native conformation and hence the potential for forming an antigenic epitope. No large differences were observed in these plots, but two smaller regions of potential difference in the hydrophilic character of VP1 were found. The region with the larger difference was from aa 63 to 75, where a moderately hydrophobic region (hydropathy index of 0.0 to 0.1) in BKV(DUN) was converted to a moderately hydrophilic region (hydropathy index of −0.9 to −1.4) in ASV. The second region was from aa 221 to 229, where a neutral region (hydropathy index of −0.4) in BKV(DUN) was converted to a weakly hydrophilic region (hydropathy index of −1.0) in ASV. The remainder of the hydropathy profiles were essentially identical.

SV40 VP1 has been reported to have a nuclear localization signal in the first 8 aa (Ala-Pro-Thr-Lys-Arg-Lys-Gly-Ser; the initial methionine is removed) (53). The ASV and BKV(DUN) VP1 proteins are identical through the first 8 aa and differ from those of SV40 only at position 8 (aa 9 in the primary translation product), with glutamic acid replacing serine.

Transformation of Rat 2 cells. The transforming activity of ASV, BKV(WT), BKV9, and BKV9(*dl*143) in Rat 2 cells was determined by a dense-focus assay (Table 2). BKV9 is a BKV variant cloned from a stock of BKV(WT); it was used as a positive control because it efficiently transforms Rat 2 cells (Table 2; 2a). The BKV9 regulatory region has been sequenced, and it differs from that of BKV(WT) only in the enhancer region (5); BKV9 contains a 143-bp tandem duplication in place of the 68–50–68-bp triplication found in BKV(WT). The transforming abilities of ASV and BKV(WT) were found to be equivalent; the foci that formed were flat and enlarged slowly. To assess the effect of reducing the number of tandem repeats of the enhancer on

TABLE 2. Transformation of Rat 2 cells by ASV and BKV

DNA[a]	No. of foci[b] (day)[c]		
	Expt 1	Expt 2	Expt 3
Calf thymus	1	1	1
pASV2	22 (22)	31 (20)	22 (24)
pBKV (WT)	11 (25)	29 (26)	31 (24)
pBKV9	163 (17)	127 (18)	172 (14)
pBKV9 (*dl*143)			126 (16)

[a] Rat 2 cells were transfected with 1.0 μg of each DNA by using the calcium phosphate coprecipitation technique (52). ASV and BKV DNAs were cloned into pBR322 at their unique *Eco*RI sites.

[b] Average number of foci per 60-mm plate (four to six plates per DNA) 30 to 38 days posttransfection.

[c] Day foci first observed. Foci were not observed on plates of cells transfected with calf thymus DNA until the cells were stained.

transforming ability, BKV9(*dl*143) was included in the third experiment. BKV9(*dl*143) is a deletion mutant of BKV9 that contains only one copy of the 143-bp duplication. Watanabe and Yoshiike (51) and Hara et al. (17) have shown that reducing the number of 68-bp repeats in BKV(WT) increases its transforming ability. BKV9(*dl*143), however, showed reduced transforming activity relative to BKV9, and the foci that formed were smaller and more compact.

Immunological characteristics. The immunological reactivity of ASV, BKV, and JCV virions to antibodies generated against all three viruses is listed in Table 3. Four antisera were used, two generated against the pool of ASV from which the clone pASV-2 was made, one against BKV(WT), and one against JCV (Mad-1 strain). HAI titers of these sera were determined by using 8 HAU of each of the viruses used to immunize the rabbits. These sera were also tested against ASV virions [ASV(cloned)] that were produced from cells transfected with pASV-2 (viral DNA removed from plasmid). The HAI titers of the four antisera were essentially identical when ASV(cloned) or the parental ASV was used as the hemagglutinating agent. Antibodies raised against BKV did not cross-react with ASV, as was found in earlier studies (7, 14). However, antibodies raised against ASV reacted equally well with BKV and ASV, in contrast to earlier studies (14), in which only a weak cross-reaction was seen between BKV and ASV by using antisera obtained 1 month postinoculation. In no case was a significant cross-reaction seen between JCV and either ASV or BKV.

From these data, it appears that ASV lacks the primary antigenic epitope(s) recognized on the BKV virion but

TABLE 3. Inhibition of hemagglutination by antibodies directed against ASV, BKV, and JCV

Immunizing agent[a]	HAI titer of viral antigens			
	ASV[b]		BKV	JCV
	(cloned)	(prototype)		
ASV	4,096[c]	8,192	8,192	128
ASV	1,024	2,048	1,024	<8
BKV	<8	<8	4,096	<8
JCV	<8	<8	<8	4,096

[a] Immunizing dose (both rabbits): ASV, 34,000 HAU; BKV (WT), 10,000 HAU; JCV (Mad 1), 64,000 HAU.

[b] ASV (cloned) indicates virus prepared from cells transfected with the molecularly cloned DNA that was separated from the plasmid sequences. ASV (prototype) indicates virus obtained from the initial isolation studies (7, 14) and from which the recombinant DNA clone was produced.

[c] HAI titers are the inverse of the greatest dilution of antibody that inhibits agglutination of human erythrocytes by 8 HAU of virus.

TABLE 4. Prevalence of antibodies against ASV, BKV, and JCV in persons in Wisconsin

Age (yr)	No. of serum samples tested	% of sera positive[a]		
		ASV	BKV	JCV
1–4	47	43	64	23
5–9	47	30	85	49
10–14	44	45	90	59
15–19	107	23	84	36
20–29	140	31	89	55
30–39	108	31	86	47
40–49	118	48	86	69
50–59	97	38	86	82
60–69	94	29	78	79
70–79	91	24	76	74
80+	84	33	80	85

[a] HAI titer of 32 or greater.

retains functional epitope(s) that are on both the ASV and BKV virions. Apparently, the immunized rabbits make a strong response to these shared epitope(s) only when the major epitope(s) of the BKV capsid is absent.

To further determine the prevalence of the human polyomaviruses in the population, 977 blood samples collected in Wisconsin were tested for the prevalence of antibodies to ASV, BKV, and JCV by the HAI assay (Table 4). Antibodies to ASV were found to be widespread in the population, with approximately 30 to 40% of the population being seropositive, in agreement with the data of Gibson and Gardner from England (14). Seroconversion usually occurs at an early age (generally by age 5), and the percentage of seropositive individuals remains fairly stable through all age groups. Exposure to BKV also occurs frequently in childhood, but the percentage of seropositives continues to increase, until 70 to 80% of the population can be shown to have circulating anti-BKV antibodies. As antibodies raised in rabbits against ASV cross-react with BKV, a portion of the sera that test positive for BKV may reflect a cross-reaction with ASV and not a true reaction with BKV. Antibodies to JCV are acquired more gradually than those to ASV or BKV, but 70 to 80% of the population eventually exhibits serological evidence of infection with JCV, as it does with BKV.

DISCUSSION

ASV was isolated from the urine of a pregnant woman in 1980 by Coleman et al. (7). Analysis by HAI and by restriction enzyme digestion (14) suggested that ASV was a new strain of human polyomavirus that was more closely related to BKV than to JCV. Serological studies indicated that ASV is endemic in the population, since 40 to 50% of the adults in England possess antibodies to the virus (14). Because ASV is incompletely characterized and is widespread in the population, we decided to define its relationship to the better-known human polyomaviruses. As a first step in the analysis, the entire ASV genome was cloned and sequenced. ASV was found to be surprisingly similar to BKV(DUN); these viruses share a homology of 94.9% at the nucleotide level.

The largest differences between ASV and BKV are found in the noncoding regulatory region. The ASV region differs considerably from those of BKV(DUN) and BKV(WT) but is similar to the regulatory regions of the BKV variants BKV(Dik) and BKV(WW) (42, 49). The arrangement of the ASV regulatory blocks is also similar to that of the evolutionary prototype, or archetype BKV, proposed by Yoshiike and Takemoto (42, 56). From the early to the late side of the regulatory region, the archetype contains a palindrome, two inverted repeats, an AT-rich region, a 68-bp block, a 39-bp block, a 63-bp block, and a putative leader region for the late message (the region is shown graphically in Fig. 2 and is represented in the text as PAL-IR-IR-TATA-68-39-63-LL). ASV differs from the archetype by lacking 31 bp at the 3′ end of the late leader plus the first base pair of the agnogene. BKV(WT) differs from the archetype more extensively than ASV; its regulatory region can be represented as PAL-IR-IR-TATA-68-50-68-39-LL. There are no direct repeats in the ASV regulatory region that are similar to the repeats found in SV40, BKV, and JCV. It has been suggested that the direct repeats found in BKV(WT) are an artifact of passage in tissue culture (42); ASV has been passed extensively in tissue culture, and it does not possess a large tandemly repeated structure within its regulatory region.

The tandem repeats found within the regulatory regions of the polyomaviruses usually represent enhancer elements that are important activators of viral transcription. Enhancers are modular collections of *cis*-acting sequence elements, each of which is thought to act by binding specific transcription factors or by altering local chromatin structure or both (for a review see reference 16). Although the ASV regulatory region does not contain a duplication, the 68-, 39-, and 63-bp blocks all contain sequences homologous to putative binding sites for transcription factors or to recognized enhancer core sequences.

Markowitz and Dynan (30) recently analyzed the binding of cellular factors to the BKV(DUN), BKV(WW), and BKV(MM) regulatory regions by DNase footprinting. These results can be extrapolated to other BKV strains because the 68-bp, 39-bp, 63-bp, and late-leader blocks that comprise these variant BKV regulatory regions are each present in at least one of the strains analyzed by these investigators. In Fig. 2 potential *cis*-acting elements are indicated for archetype BKV, ASV, BKV(WT), and BKV9. Elements included are binding sites for NF-BK, Sp1, and AP1. Also shown is a binding site (termed L1) for an unknown factor that was detected by these investigators. NF-BK is a member of the nuclear factor I family (9); it is provisionally referred to as NF-BK until its relationship to other members of the family can be determined. The consensus NF-BK binding site is TGGAA(T/A)(G/C)(C/T)(A/T)GCCAAA (30). The transcriptional activator Sp1 binds the consensus sequence (G/T)(G/A)GGCG(G/T)(G/A)(G/A)(C/T) (15). AP1 binds to the sequence TGACTCA (27). Potential AP1 binding sites were not identified from ASV sequence information, although sites are predicted in BKV(WT). ASV does not contain the L1 site (it was located in the 32 bp lost from the late-leader region); therefore, binding to L1 is not essential for viral growth. The L1 binding site is found in BKV(WT).

Deyerle and co-workers (8) recently analyzed the BKV(WT) early promoter and enhancer. By their analysis, the minimal BKV early promoter is composed of the sequences PAL-IR-IR-TATA plus the early portion of the first 68-bp block. ASV contains the same arrangement, and these sequences probably form the ASV early promoter. Deyerle and co-workers found that the optimal BKV(WT) enhancer consists of two 68-bp blocks and an element termed C (located in the 39-bp block, overlapping an NF-BK site). The ASV regulatory region contains one copy of the 68-bp block plus the C element; it also contains a 63-bp block. The 63-bp block in ASV might substitute for a 68-bp block of BKV(WT) to yield full ASV enhancer activity.

A deletion mutant of BKV(WT) containing only one 68-bp block (*dl*504, PAL-IR-IR-TATA-68-C-LL) demonstrates an increased transforming ability (17, 51). ASV contains only one copy of the 68-bp block, yet a significant difference in the transforming abilities of ASV and BKV(WT) was not observed. If the 63-bp element can substitute for a 68-bp element, ASV would resemble the BKV(WT) deletion mutant *dl*503 (PAL-IR-IR-TATA-68-68-C-LL). *dl*503, like BKV(WT), transforms cells inefficiently (51). BKV9 contains a tandem duplication of 143 bp in the enhancer region, yet it transforms efficiently. Inspection of the BKV9 regulatory region shows that, although it includes approximately the same number of potentially active transcription signals as does BKV(WT) (Fig. 2), they are arranged in two groups separated by a short length of the late coding region. This novel arrangement may alter the interaction of potential regulatory proteins that bind to these sites. The cumulative effects of this arrangement are unknown, but they may be sufficient to produce an elevated transforming activity similar to that found in *dl*504 (51). The regulatory region of BKV9(*dl*143) is nearly identical to that of *dl*504; this virus also transforms Rat 2 cells efficiently.

The transforming T antigens of ASV and BKV(DUN) are nearly identical, but the extreme carboxy-terminal region of the ASV T antigen lacks 4 aa found in the corresponding region of the BKV(DUN) T antigen [BKV(DUN) aa 664-665 deleted and aa 668-671 deleted and replaced by Glu-Asn]. Alterations to the analogous sequences of SV40 T antigen (aa 675 to 681) lead to reduced virus yields and cold sensitivity for growth in the CV1-P and BSC cell lines (6, 39). The adenovirus helper function of SV40 T antigen has been mapped to aa 674 to 699 (6), and this function in SV40 may contribute to posttranscriptional regulation of the agnoprotein (21). When grown in CV1-P or BSC cells, mutant viruses lacking SV40 T antigen aa 671 to 708 (*dl*1066) or aa 676 to 685 (*dl*1140) express the late mRNAs and the late structural proteins at a 5- to 15-fold lower level than does wild-type virus, but they express the agnoprotein at a level that is 100-fold lower. If an analogous posttranscriptional mechanism operates in ASV and BKV(DUN), the deletions in ASV in the T antigen and in the late-leader region upstream of the agnogene could disrupt this mechanism and alter the levels of agnoprotein produced in infection.

The deletion of 32 bp in the late-leader region of ASV alters the agnogene initiation codon and shifts a potential replacement initiation codon 24 bp upstream. If this alternate ATG is functional, the proposed agnoprotein would be increased in size from 66 aa in BKV(DUN) to 74 aa in ASV. Changes to the coding sequence would be limited to the addition of the sequence Met-Phe-Cys-Glu-Pro-Lys-Asn-Leu at the amino terminus and replacement of the original amino-terminal methionine with an internal valine. Another difference between the ASV and BKV(DUN) agnoproteins might be their levels of expression; the translational efficiency of the new start site may differ from that of the original site. It is also possible that alteration of the start site prevents synthesis of the ASV agnoprotein or yields a nonfunctional protein. This would not necessarily be inconsistent with the finding that ASV is viable, since the SV40 agnoprotein has been shown to be nonessential for growth in vitro (32, 48). Resnick and Shenk (41) have shown that mutating the SV40 agnogene ATG initiation codon to TTG (*pm*1493, SV40 335 A to T) prevents agnoprotein synthesis, yet the mutant virus is viable in tissue culture.

DNA sequence homology data support the suggestion that the alternate ATG of the ASV agnoprotein is functional and that this protein is expressed. Beyond the amino-terminal sequences, the ASV and BKV(DUN) agnoprotein sequences have been conserved; there are no insertions or deletions within this 225-bp gene, indicating that it has been maintained through selective pressure in vivo. In contrast, all of the predicted noncoding regions in ASV contain insertions or deletions when compared to BKV(DUN). Inspection of the sequences surrounding the putative ASV agnogene initiation codon yields additional evidence that this signal is utilized. The consensus flanking sequence for eucaryotic mRNA initiation codons is CC(A/G)CCAUGG (22, 23). The flanking sequence for the BKV(DUN) agnogene initiation codon is AGGCCAUGG; for the putative ASV agnogene initiation codon it is CAGACAUGT. Kozak (22) inspected 211 mRNA leader sequences and found that a purine was present at position −3 (A of AUG is +1) in 97% of these messages. Although A is favored in the −3 position, G is frequently found. Because the late mRNAs that code for the agnoprotein are polycistronic, it may be necessary that the consensus flanking sequences of the BKV(DUN) and ASV agnogene start sites be imperfect. The agnoprotein message is located within the common 5′ domain of the late mRNA and hence is present in both the 19S VP2 and VP3 and in the 16S VP1 mRNAs. Flanking sequences of the initiation codons affect the relative levels of synthesis from the various open reading frames in a polycistronic eucaryotic mRNA (24, 25). Thus, the nearly equivalent initiation flanking sequences for the ASV and BKV(DUN) agnogenes may indicate a similar level of coordination of late protein translation in the two viruses.

The distinct antigenicity of ASV is intriguing given the high degree of homology between the structural proteins of ASV and BKV(DUN). A possible explanation for the differing immunogenicity of these two viruses is based on the 38-aa differences in their late structural proteins. These amino acid changes produce two noticeable differences in the hydrophobicity profiles of the ASV and BKV(DUN) VP1 proteins; one of these differences represents a distinct cluster of altered amino acid residues in the amino portion of VP1. It is possible that the point mutations have removed the epitopes recognized on the BKV virion, leaving only weaker epitopes that are not normally recognized in BKV to be recognized on the ASV virion. As these weaker epitopes are found on the BKV virion, the antibodies generated by ASV would also react with BKV.

An alternative explanation for the distinct antigenicity of ASV involves the interaction of VP1 with the agnoprotein. The function of the agnoprotein in polyomaviruses is not well understood, but a number of activities have been ascribed to the SV40 protein. The SV40 agnoprotein participates in virion production after the removal of DNA from the replication pathway (20); it binds VP1 in infected cells and helps to temporally regulate the nuclear localization of VP1 (3). In cells infected with agnogene deletion mutants, virion burst size is decreased 17- to 100-fold (33, 41), but the kinetic pattern of VP1 polymerization is unchanged; it remains cooperative (33). An activity of agnoprotein may also be involved in host cell lysis (41). The SV40 agnoprotein may act by inhibiting VP1 polymerization until the latter protein can productively encapsidate minichromosomes in the nucleus (1). On the basis of the activities of the SV40 agnoprotein, a number of possibilities can be proposed to explain how the altered ASV agnoprotein might affect the antigenicity of the capsid. While bound to VP1 in the cytoplasm, the altered amino terminus of the ASV agnoprotein may affect the posttranslational modification of the major capsid protein and hence modify an epitope. Alternatively, the ASV agnoprotein may affect the polymerization pattern of VP1 through its interaction with this protein in the cytoplasm and nucleus. In mouse polyomavirus, purified VP1 can polymerize in vitro into a number of capsidlike structures (43) and hence may alter the availability of different antigenic epitopes to be recognized. These possibilities for the generation of altered antigenicity of ASV relative to BKV need not be independent or exhaustive.

A scheme for the divergence of ASV from the archetype BKV can be proposed that rests on two assumptions: (i) an ASV agnoprotein is produced which functions in a manner that is similar to that of its SV40 counterpart, and (ii) posttranscriptional regulation of expression of the human polyomavirus agnoproteins occurs via T antigen. The initial event in the divergence of ASV from the archetype might have been the deletion of sequences from the late-leader region, thereby altering the agnogene product and disrupting its normal regulation by T antigen. Secondary events would then have involved the alteration of the T antigen and VP1 coding sequences to adapt to this primary mutation. Deletion of sequences in the carboxy-terminal domain of T antigen might have allowed ASV to reestablish proper regulation of the translation of the agnogene. The occurrence of a number of point mutations within the VP1 coding region may have been important to the maintenance of a productive interaction between VP1 and agnoprotein (second-site revertants are found in SV40 in the agnoprotein when VP1 is mutated [29] and in VP1 when the agnoprotein is mutated [1]). In turn, these mutations in VP1 may have resulted in an alteration of the primary antigenic epitopes of the ASV capsid, removing the epitopes recognized in the archetype BKV capsid. This would have given ASV a selective advantage in vivo, since the virus would not have been inactivated by anti-BKV antibodies; ASV would spread through the population without pressure from previous exposure of the host to BKV.

The sequence data presented here indicate that ASV is not a distinct human polyomavirus but rather an antigenic variant of BKV. We therefore suggest that its name be changed from ASV to BKV(AS) to better reflect this relationship.

ACKNOWLEDGMENTS

We thank R.-B. Markowitz and W. S. Dynan for access to their data prior to publication and for helpful comments on the manuscript and Dawn Wooley for construction of BKV9(*dl*143).

This work was supported by Public Health Service grants CA-44970 and AI11217 from the National Cancer Institute and the National Institute of Allergy and Infectious Diseases, respectively, and by a grant from the Tracy Ruhrup Memorial Fund. J.T. was supported by a National Science Foundation Graduate Fellowship.

LITERATURE CITED

1. **Barkan, A., R. C. Welch, and J. E. Mertz.** 1987. Missense mutations in the VP1 gene of simian virus 40 that compensate for defects caused by deletions in the viral agnogene. J. Virol. **61**:3190–3198.
2. **Birnboim, C., and J. Doly.** 1979. Rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.

2a. **Bollag, B., W.-F. Chuke, and R. J. Frisque.** 1989. Hybrid genomes of the polyomaviruses JC virus, BK virus, and simian virus 40: identification of sequences important for efficient transformation. J. Virol. **63**:863–872.

3. **Carswell, S., and J. C. Alwine.** 1986. Simian virus 40 agnoprotein facilitates perinuclear-nuclear localization of VP1, the major capsid protein. J. Virol. **60**:1055–1061.

4. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:**165–170.
5. **Chuke, W.-F., D. L. Walker, L. B. Peitzman, and R. J. Frisque.** 1986. Construction and characterization of hybrid polyomavirus genomes. J. Virol. **60:**960–971.
6. **Cole, C. N., and T. P. Stacy.** 1987. Biological properties of simian virus 40 host range mutants lacking the COOH-terminus of large T antigen. Virology **161:**170–180.
7. **Coleman, D. V., M. R. Wolfendale, R. A. Daniel, N. K. Dhanjal, S. D. Gardner, P. E. Gibson, and A. M. Field.** 1980. A prospective study of human polyomavirus infection in pregnancy. J. Infect. Dis. **142:**1–8.
8. **Deyerle, K. L., J. A. Cassill, and S. Subramani.** 1987. Analysis of the early regulatory region of the human papovavirus BK. Virology **158:**181–193.
9. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell **50:**863–872.
10. **Fiers, W., R. Contreras, G. Haegerman, R. Rogires, A. van der Voorde, H. van Heuverswyn, J. van Herreweghe, B. Volckaert, and M. Ysebaert.** 1978. Complete nucleotide sequence of SV40 DNA. Nature (London) **273:**113–120.
11. **Frisque, R. J., G. L. Bream, and M. T. Cannella.** 1984. Human polyomavirus JC virus genome. J. Virol. **51:**458–469.
12. **Gardner, S. D.** 1973. Prevalence in England of antibody to human polyomavirus (BK). Br. Med. J. **1:**77–78.
13. **Gardner, S. D., A. M. Field, D. V. Coleman, and B. Hulme.** 1971. New human papovavirus (BK) isolated from urine after renal transplantation. Lancet **i:**1253–1257.
14. **Gibson, P. E., and S. D. Gardner.** 1983. Strain differences and some serological observations on several isolates of human polyomaviruses, p. 119–132. *In* J. L. Sever and D. L. Madden (ed.), Polyomaviruses and human neurological diseases. Alan R. Liss, Inc., New York.
15. **Gidoni, D., J. T. Kadonaga, H. Barrera-Saldana, K. Takahashi, P. Chambon, and R. Tjian.** 1985. Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science **230:**511–517.
16. **Gluzman, Y.** 1985. Eukaryotic transcription. The role of cis- and trans-acting elements in initiation. *In* Y. Gluzman (ed.), Current communications in molecular biology. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Hara, K., Y. Oya, and Y. Yogo.** 1985. Enhancement of the transforming capacity of BK virus by partial deletion of the 68-base-pair tandem repeats. J. Virol. **55:**867–869.
18. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
19. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
20. **Hou-Jong, M.-H. J., S. H. Larsen, and A. Roman.** 1987. Role of the agnoprotein in regulation of simian virus 40 replication and maturation pathways. J. Virol. **61:**937–939.
21. **Khalili, K., J. Brady, J. M. Pipas, S. L. Spence, M. Sadofsky, and G. Khoury.** 1988. Carboxyl-terminal mutants of the large tumor antigen of simian virus 40: a role for the early protein late in the lytic cycle. Proc. Natl. Acad. Sci. USA **85:**354–358.
22. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12:**857–872.
23. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:**283–292.
24. **Kozak, M.** 1986. Bifunctional messenger RNAs in eukaryotes. Cell **47:**481–483.
25. **Kozak, M.** 1986. Regulation of protein synthesis in virus-infected animal cells, p. 229–292. *In* K. Maramorosch, F. A. Murphy, and A. J. Shatkin (ed.), Advances in virus research, vol. 31. Academic Press, Inc., Orlando, Fla.
26. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
27. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell **49:**741–752.
28. **Manaker, R. A., G. Khoury, and C.-J. Lai.** 1979. The spliced structure of BK virus mRNAs in lytically infected and transformed cells. Virology **97:**112–121.
29. **Margolskee, R. F., and D. Nathans.** 1983. Suppression of a VP1 mutant of simian virus 40 by missense mutations in serine codons of the viral agnogene. J. Virol. **48:**405–409.
30. **Markowitz, R.-B., and W. S. Dynan.** 1988. Binding of cellular proteins to the regulatory region of BK virus DNA. J. Virol. **62:**3388–3398.
31. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
32. **Mertz, J. E., and P. Berg.** 1974. Viable deletion mutants of simian virus 40: selective isolation by means of restriction endonuclease from *Hemophilus parainfluenzae*. Proc. Natl. Acad. Sci. USA **71:**4879–4883.
33. **Ng, S.-C., J. E. Mertz, S. Sanden-Will, and M. Bina.** 1985. Simian virus 40 maturation in cells harboring mutants deleted in the agnogene. J. Biol. Chem. **260:**1127–1132.
34. **Padgett, B. L., C. M. Rogers, and D. L. Walker.** 1977. JC virus, a human polyomavirus associated with progressive multifocal leukoencephalopathy: additional biological characteristics and antigenic relationships. Infect. Immun. **15:**656–662.
35. **Padgett, B. L., and D. L. Walker.** 1973. Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leukoencephalopathy. J. Infect. Dis. **127:**467–470.
36. **Padgett, B. L., D. L. Walker, G. M. Zu Rhein, R. J. Eckroade, and B. H. Dessel.** 1971. Cultivation of papovalike virus from human brain with progressive multifocal leucoencephalopathy. Lancet **i:**1257–1260.
37. **Pagnani, M., M. Negrini, P. Reschiglian, A. Corallini, P. G. Balboni, S. Scherneck, G. Macino, B. 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.
38. **Pater, A., M. M. Pater, L.-S. Chang, K. Slawin, and G. I. 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.
39. **Pipas, J. M.** 1985. Mutations near the carboxyl terminus of the simian virus 40 large tumor antigen alter viral host range. J. Virol. **54:**569–575.
40. **Proudfoot, N. J., and G. G. Brownlee.** 1974. Sequence at the 3′ end of globin mRNA shows homology with immunoglobulin light chain mRNA. Nature (London) **252:**359–362.
41. **Resnick, J., and T. Shenk.** 1986. Simian virus 40 agnoprotein facilitates normal nuclear location of the major capsid polypeptide and cell-to-cell spread of virus. J. Virol. **60:**1098–1106.
42. **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.
43. **Salunke, D. M., D. L. D. Caspar, and R. L. Garcea.** 1986. Self-assembly of purified polyomavirus capsid protein VP1. Cell **46:**895–904.
44. **Seif, I., G. Khoury, and R. Dhar.** 1979. The genome of human papovavirus BKV. Cell **18:**963–977.
45. **Seif, I., G. Khoury, and R. Dhar.** 1981. Sequence and analysis of the genome of human papovavirus BKV, p. 911–936. *In* J. Tooze (ed.), Molecular biology of tumor viruses, part 2, 2nd ed. DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
46. **Sever, J. L.** 1962. Application of a microtechnique to viral serological investigations. J. Immunol. **88:**320–329.
47. **Sompayrac, L. M., and K. J. Danna.** 1981. Efficient infection of monkey cells with DNA of simian virus 40. Proc. Natl. Acad. Sci. USA **78:**7575–7578.
48. **Subramanian, K. N.** 1979. Segments of simian virus 40 DNA

spanning most of the leader sequence of the major late viral messenger RNA are dispensable. Proc. Natl. Acad. Sci. USA **76**:2556–2560.

49. **ter Scheggett, 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.
50. **Topp, W. C.** 1981. Normal rat cell lines deficient in nuclear thymidine kinase. Virology **113**:408–411.
51. **Watanabe, S., and K. Yoshiike.** 1985. Decreasing the number of 68-base-pair tandem repeats in the BK virus transcriptional control region reduces plaque size and enhances transforming capacity. J. Virol. **55**:823–825.
52. **Wigler, M., A. Pellicer, S. Silverstein, and R. Axel.** 1978. Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell **14**:725–731.
53. **Wychowski, C., D. Benichou, and M. Girard.** 1986. A domain of SV40 polypeptide VP1 that specifies migration into the cell nucleus. EMBO J. **5**:2569–2576.
54. **Wychowski, C., D. Benichou, and M. Girard.** 1987. The intranuclear location of simian virus 40 polypeptides VP2 and VP3 depends on a specific amino acid sequence. J. Virol. **61**:3862–3869.
55. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. Science **206**:456–462.
56. **Yoshiike, K., and K. K. Takemoto.** 1986. Studies with BK virus and monkey lymphotropic papovavirus. *In* N. Salzman (ed.), The papovaviridae, vol. 1. The polyomaviruses. Plenum Publishing Corp., New York.