Journal of General Virology (2004), **85**, 2651–2663 **DOI** 10.1099/vir.0.79920-0

Analysis of 15 novel full-length BK virus sequences from three individuals: evidence of a high intra-strain genetic diversity

Yiping Chen,[1] Paul M. Sharp,[2] Mary Fowkes,[3] Olivier Kocher,[3] Jeffrey T. Joseph[4] and Igor J. Koralnik[1,4]

[1,3,4]Division of Viral Pathogenesis[1] and Departments of Pathology[3] and Neurology[4], Beth Israel Deaconess Medical Center, Harvard Medical School, RE 213B, 330 Brookline Avenue, Boston, MA 02215, USA

[2]Institute of Genetics, University of Nottingham, Queens Medical Centre, Nottingham, UK

Correspondence
Igor J. Koralnik
ikoralni@bidmc.harvard.edu

Received 18 December 2003
Accepted 11 May 2004

To determine the variability of BK virus (BKV) *in vivo*, the sequences of nine full-length molecular clones from the striated muscle and heart DNA of a patient with BKV-associated capillary leak syndrome (BKV_{CAP}), as well as three clones each from the urine of one human immunodeficiency virus type 2-positive (BKV_{HI}) and one healthy control subject (BKV_{HC}), were analysed. The regulatory region of all clones corresponded to the archetypal regulatory region usually found in urine isolates. Analysis of the predicted conformation of BKV_{CAP} proteins did not suggest any structural differences on the surface of the viral particles compared with BKV_{HI} and BKV_{HC} clones. No amino acid changes common to most BKV_{CAP} clones could be identified that have not already been reported in non-vasculotropic strains. However, the coding region of each clone had unique nucleotide substitutions, and intra-host variability was greater among BKV_{CAP} clones, with a mean difference of 0·29 % per site compared with 0·16 % for BKV_{HI} and 0·14 % for BKV_{HC}. The clones from each strain formed monophyletic clades, suggesting a single source of infection for each subject. The most divergent BKV_{CAP} clones differed at 0·55 % of sites, implying a rate of nucleotide substitution of approximately 5×10^{-5} substitutions per site per year, which is two orders of magnitude faster than estimated for the other human polyomavirus, JC virus.

INTRODUCTION

There are two known human polyomaviruses, JC virus (JCV) (Padgett *et al.*, 1971) and BK virus (BKV) (Gardner *et al.*, 1971). Asymptomatic infection with these polyomaviruses occurs in childhood in more than 80 % of healthy individuals (Knowles *et al.*, 1989; Tavis *et al.*, 1990; Weber *et al.*, 1997). JCV is the aetiological agent of progressive multifocal leukoencephalopathy (PML) (Berger *et al.*, 1987), which occurs in patients with AIDS or haemopathies and in organ transplant recipients. In bone marrow recipients, BKV can cause haemorrhagic cystitis (Rice *et al.*, 1985), and in renal transplant recipients, it has been associated with ureteric stenosis (Coleman *et al.*, 1978) and with a severe tubulointerstitial nephritis, which may lead to loss of the allograft (Hirsch *et al.*, 2002; Nickeleit *et al.*, 1999).

Since PML was identified as a major opportunistic infection at the beginning of the AIDS epidemic, JCV pathogenicity has been widely studied and full-length sequences of more than 250 strains of this virus have been analysed. Conversely, BKV has only been recognized as an important human pathogen more recently and until now the full-length sequences of only three strains have been reported. BKV strains have been classified into four different genotypes, which correspond to distinct serotypic profiles on the basis of nucleotide changes within a small fragment of the VP1 gene.

Little is know about polyomavirus sequence variation within individuals. There have been a few reports showing the co-existence of multiple JCV-coding region sequences in a single patient (Agostini *et al.*, 1996; Ferrante *et al.*, 2001; Martin & Foster, 1984). BKV sequence variations were observed on sequential kidney biopsy specimens from patients with polyomavirus nephropathy (Randhawa *et al.*, 2002). However, it has been suggested that polyomaviruses have co-evolved with their hosts (Shadan & Villarreal, 1993; Soeda *et al.*, 1980) and, consistent with this, the major genotypes of JCV appear to have diverged in parallel with ancient human migrations (Agostini *et al.*, 1997; Guo *et al.*,

The GenBank/EMBL/DDBJ accession numbers for the nucleotide sequences determined in this work are AY628224–AY628238.

1998). Under the assumption of host–virus co-evolution, the rate of evolution of polyomaviruses has been estimated to be in the range of 4×10^{-7} to 4×10^{-8} synonymous nucleotide substitutions per site per year (Hatwell & Sharp, 2000; Yasunaga & Miyata, 1982). These rates are so slow as to suggest that almost no genetic polymorphisms would be expected to be found among the viral sequences within a given individual, unless this person had been infected by more than one strain. Thus, a more specific analysis of intra-host genetic diversity of BKV is needed.

We have recently reported the case of a renal transplant recipient infected with a novel BK-related virus that had a tropism for capillary endothelial cells (Petrogiannis-Haliotis *et al.*, 2001). This infection resulted in a vasculopathy that led to a capillary leak syndrome, myocardial infarction and death. To determine whether this BKV isolate was different from the known BKV strains and to explore BKV variability *in vivo*, we have sequenced nine full-length clones of this virus amplified from the heart and striated muscle DNA of this patient. Since the three published full-length BKV sequences have been derived from isolates that had been cultured *in vitro*, which may promote artificial mutations and lead to the selection of laboratory-adapted strains, we characterized the sequences of six novel full-length BKV clones amplified from the urine of one human immunodeficiency virus type 2 (HIV-2)-positive and one healthy individual and used them as additional controls. Our results indicated that a surprising amount of BKV sequence variation exists within a given individual. This diversity sheds new light on our understanding of BKV evolution.

METHODS

Specimen collection and nucleic acid extraction. Fresh-frozen autopsy samples of heart and striated muscle from the patient with BKV vasculopathy (Petrogiannis-Haliotis *et al.*, 2001) were kept at −80 °C until DNA extraction. Tissues were cryo-sectioned and 100 mg blocks were pulverized using a pestle and mortar, resuspended in TNE buffer (10 mM Tris/HCl, 100 mM NaCl, 10 mM EDTA), lysed by adding SDS (1 % final concentration) and incubated overnight in the presence of proteinase K at 37 °C. DNA was extracted with phenol/chloroform/isoamyl alcohol and precipitated with ethanol. Twenty millilitre urine samples were collected from 15 HIV-positive and 30 HIV-negative patients and stored at −20 °C until DNA extraction. Urine sediments were collected by centrifugation at 25 000 ***g***, resuspended in TNE buffer and DNA was extracted as described above.

Qualitative BKV PCR. The presence of BKV DNA in the clinical samples was determined by PCR. A total of 200 ng DNA extracted from tissue or urine samples was used in a 50 μl reaction consisting of Platinum PCR Supermix (Invitrogen) and 10 pmol of each of the oligonucleotide primers BK2116F and BK2415R (Table 1), which amplified a 300 bp fragment of the BKV VP1 gene, or JCV-specific primers VP11 and VP12 (Koralnik *et al.*, 1999), or SV40-specific primers SP15 5′-TTAGCAGCTGAAAAACAGTTTACAGAT-3′ (nt 1712–1738) and SP13 5′-TTAACAGTAACAGCTTCCCACATCA-3′ (nt 1830–1854), which amplified a 181 bp fragment of the JCV and a 143 bp fragment of the SV40 VP1 genes, respectively. The amplification was carried out in a GeneAmp PCR System 9700 (PE Applied Biosystems) with a first denaturing step of 2 min at 94 °C, followed by 40 cycles consisting of 30 s denaturation at 94 °C, 30 s annealing at 55 °C and 1 min elongation at 72 °C, with a final elongation period of 7 min at 72 °C. Amplified products were analysed by electrophoresis on a 1 % agarose gel.

Long BKV PCR, molecular cloning and DNA sequencing. The *Taq* Plus Long PCR system (Stratagene) was used to amplify full-length BKV genomes from clinical samples in one reaction. One microgram of DNA extracted from fresh-frozen autopsy samples from the heart and striated muscle of the BKV vasculopathy patient, as well as from the urine samples of one HIV-2-positive patient and one HIV-negative healthy control subject, was digested with *Bam*HI and added to a 50 μl reaction mixture containing 25 pmol of each of the oligonucleotide primers BK1731F and BK1739R (Table 1). These inverse PCR primers overlapped at the *Bam*HI site of the BKV genome and were used to amplify the full-length linearized BKV genome. The PCR was performed with a first denaturing step of 2 min at 94 °C, followed by 10 cycles consisting of 30 s denaturation at 94 °C, 1 min annealing at 55 °C and 5·5 min elongation at 68 °C, followed by 30 cycles of 30 s denaturation at 94 °C, 1 min annealing at 65 °C and 10 min elongation at 68 °C, with a final elongation period of 30 min at 68 °C. The 5 kb PCR products were separated by gel electrophoresis, cut from the gel with a sterile disposable scalpel, purified using the QIAquick Gel Extraction kit and cloned into the pCR2.1-TOPO or pCR-XL-TOPO vector. Single *Escherichia coli* colonies (Top 10 strain) were picked and the plasmids purified. Prior to sequencing the full-length BKV clones, the integrity of the *Bam*HI site was verified by PCR using 10 pmol each of primers BK1319F and BK1811R (Table 1) in a 50 μl reaction containing the Platinum PCR Supermix reagent and 100 ng template plasmid. A 493 bp PCR product was purified from the agarose gel, cloned into the pCR2.1-TOPO vector and sequenced. Clones that contained an intact *Bam*HI site were selected for sequencing of the full-length genome, using both forward and reverse oligonucleotide primers as listed in Table 1.

Determination of the error rate of the *Taq* Plus Long PCR system. To determine the error rate of the long PCR technique used for the BKV_{CAP}, BKV_{HI} and BKV_{HC} strains, the *Taq* Plus Long PCR system was used to amplify the BKV Dunlop reference strain (BKV_{DUN}) obtained from the American Type Culture Collection (ATCC). The pBR322-BKV_{DUN} plasmid (containing the full-length BKV_{DUN} DNA inserted at the *Bam*HI site) was digested with *Bam*HI to separate the BKV genome from the plasmid and with *Ava*I and *Pvu*I to degrade the vector for 2 h at 37 °C in a water bath. After inactivation of the restriction enzymes at 70 °C for 30 min, 25 ng BKV_{DUN} genome DNA and 1 μg human genome DNA were added to each PCR. The long PCR conditions were as described above. The 5 kb PCR-amplified products from 10 individual PCRs were purified from 0·5 % agarose gel using the QIAquick Gel Extraction Kit (Qiagen) and cloned into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen). A 1360 bp (nt 2154–3513) fragment of one clone from each PCR was sequenced on both strands with sequencing primers BK2723F and BK3184R.

Laser capture microdissection (LCM). This was performed on a fresh-frozen autopsy sample of the heart from the patient with BKV vasculopathy. Frozen tissues were cryo-sectioned on a Laser Pressure Catapulting 1·35 μm thin polyethylene film (PALM Microlaser Technologies), fixed in 70 % ethanol for 5 min and stained with 0·5 % toluidine blue in water.

Enlarged capillary endothelial cells containing nuclear inclusions or myocytes that appeared normal were captured using the Microbeam Laser apparatus (PALM Microlaser Technologies). One to six cells were captured in each microdissection. DNA from approximately 100 cells was extracted using the Gentra Systems Puregene DNA Isolation kit and resuspended in 40 μl Gentra DNA hydration solution. PCR

Table 1. Oligonucleotide primers used for PCR and sequencing of the BKV genome

Primer*	Sequence (5′→3′)
BK345F	CTAGGAATCTTGGCCTTGTCCCC
BK831F	GTAATTGCTGGTGCTCCTGGGGC
BK1319F	AGTAGCTGAAAGGGAAGGTACCCG
BK1731F	GGGGGATCCAGATGAAAACCTTAGGGGCTTTAG
BK2116F	GCCCAGTCCCAGGTAATGAATA
BK2217F	TGCTAGGTATTTTGGGACTTTCACA
BK2239F	ACAGGAGGGGAAAATGTTCCCCC
BK2723F	TATTTTGGGGGTGGTGTTTTAGGCC
BK3216F	CATTCATTGTAACCAAGCCTGGTGG
BK3712F	TCTTTCTGTTAGCATTTCTTCCCTGG
BK4075F	GTATGGTATGGATCTCTAGTTAAGGC
BK4857F	AGGCCATTCCTTGCAGTACAGGG
BK5106F	AAGGTCCATGAGCTCCATGGATTCTTCC
BK128R	AATATATAAGAGGCCGAGGCCGCC
BK657R	CTAGGTCCCCCAAAAGTGCTAGAGCAGC
BK740R	TTCTATAGCAGCAGCAGCCTCCCC
BK1223R	AGTTTCCTCCAAAAATCTAGCCAAGG
BK1739R	GGATCCCCCATTTCTGGGTTTAGGAAGCATTCTAC
BK1811R	CTCTCTGGGCTATCACTGCTAAAGTCATTTTCAGC
BK2213R	TCATTTCTACTGGGATCAGGTACCC
BK2320R	GCCCCACACCCTGTTCATC
BK2329R	GCCCCACACCCTGTTCATC
BK2415R	TGCAAGGCCTCTCCACTGTT
BK2715R	CTTAAAAGTGGCTTATACAAAAGCAGC
BK3184R	ATTTGTAAGACAAATAGATTTTAGGCCC
BK3643R	TATGGCAGGTGTTGCTTGGCTGC
BK4076R	CTATAGAAGAAAGCATTCAAGGGGGC
BK4461R	GAAGCAACAGCAGATTCTCAACACTCAACACCACCC

*The numbers in the primer names correspond to the position of the primer 5′ end based on the BKV_{DUN} sequence. F, Forward primer; R, reverse primer.

amplification was performed using 10 pmol of the BK2217F and BK2320R primers (Table 1), which amplified a 104 bp fragment of the BKV VP1 gene, or the VP11 and VP12 primers, which amplified a 181 bp fragment of the JCV VP1 gene (Koralnik *et al.*, 1999), in a 50 µl reaction volume containing the Platinum PCR Supermix and 10 µl DNA solution. Amplified products were subjected to direct sequencing.

Evolutionary analysis of BKV sequences. Previously published BKV sequences were obtained from GenBank (Table 2). DNA sequences were aligned and compared using CLUSTAL W (Thompson *et al.*, 1994). Maximum-likelihood phylogenetic analyses were performed using DNAML from the PHYLIP package (Felsenstein, 1993). The input order of sequences was shuffled in 10 replicate analyses and the transition-to-transversion ratio was optimized at 1·9. The extent of synonymous and non-synonymous nucleotide substitution within coding sequences was estimated by the method of Li (1993) and compared with previous estimates for JCV (Hatwell & Sharp, 2000).

Conformational analysis of the VP1 capsid protein. Conformational analysis of the VP1 protein of the BKV_{CAP} clones was performed by comparison with the structure of SV40 VP1 pentamers using the O display software (Jones *et al.*, 1991) with the kind assistance of Dr Stephen Harrison.

RESULTS

Qualitative PCR

DNA is usually better preserved in fresh-frozen biopsy samples than in autopsy specimens, where tissues have already undergone autolysis and partial degradation of nucleic acids. Since our goal was to obtain a full-length genome of the vasculopathy-associated polyomavirus, we initially attempted to amplify polyomavirus sequences from the muscle biopsy sample of this patient (Petrogiannis-Haliotis *et al.*, 2001). However, numerous attempts yielded only a small 136 bp fragment of the T gene with 96 % sequence identity to BKV_{DUN} (Seif *et al.*, 1979). When DNA extracted from fresh-frozen autopsy samples for the heart or muscle of this patient was used as template, polyomavirus sequences were readily detectable with BKV, but not JCV- or SV40-specific oligonucleotide primers (data not shown). Polyomavirus particles had previously been demonstrated in capillary endothelial cells of this patient by electron microscopy (Petrogiannis-Haliotis *et al.*, 2001).

Table 2. Origin of BKV samples

BKV strains	Diagnosis	Sample type	Genetic type	Geographical origin	Ethnicity†	Method of isolation‡	Reference for sequence data
Dun	Renal transplant	Urine	I	Sudan	NA	VC+MC	Seif *et al.* (1979)
MM	Wiskott–Aldrich syndrome	Urine and brain	I	USA	NA	VC+MC	Yang & Wu (1979)
AS	Pregnant woman	Urine	III	UK	NA	VC+MC	Tavis *et al.* (1989)
JL*	Bone marrow	Urine	I	The Netherlands	NA	VC+MC	Sugimoto *et al.* (1990)
DIK*	Acute tonsillitis	Urine	I	The Netherlands	NA	VC+MC	Sugimoto *et al.* (1990)
MT*	Systemic lupus erythematosus	Urine	I	Japan	NA	MC	Sugimoto *et al.* (1990)
WW*	Renal transplant	Urine	I	South Africa	NA	MC	Sugimoto *et al.* (1990)
SB*	Lymphoma	Urine	II	UK	NA	VC+MC	Jin *et al.* (1993a)
CAP	Capillary leak syndrome	Muscle and heart	I	USA	C	PCR+MC	This study
HI	HIV-2 infection	Urine	I	West Africa	AC	PCR+MC	This study
HC	Healthy control	Urine	I	USA	AA	PCR+MC	This study

*Partial sequence available (JL, Dik and MT: agnoprotein, VP1, VP2 and VP3 genes; WW: regulatory region and agnoprotein, VP1, VP2 and VP3 genes; SB: VP1 gene).

†NA, Not available; C, Caucasian; AC, African–Creole; AA, African–American.

‡VC, Viral culture; MC, molecular cloning.

Detection of BKV DNA in capillary endothelial cells after LCM

To confirm that these polyomavirions were BKV, capillary endothelial cells harbouring enlarged nuclei were sampled by LCM and their DNA was extracted and subjected to PCR amplification. A BKV, but not JCV, VP1 gene fragment was successfully amplified from capillary endothelial cells, but not from myocytes collected by LCM (data not shown), which confirmed that BKV was indeed located exclusively in the capillary endothelial cells of this patient (Petrogiannis-Haliotis *et al.*, 2001).

Only three BKV strains, DUN (Gardner *et al.*, 1971; Seif *et al.*, 1979), MM (Takemoto *et al.*, 1974; Yang & Wu, 1979) and AS (Coleman *et al.*, 1980; Tavis *et al.*, 1989), have been entirely sequenced. These and most partially sequenced BKV strains published to date have been derived from isolates that have been cultured *in vitro* (Table 2), which may induce mutations in their DNA, especially at the level of their non-coding regulatory region (Johnsen *et al.*, 1995). Since our goal was to compare the entire genome of this BKV variant with other strains of BKV *ex vivo*, we screened urine samples from 45 individuals (15 HIV-positive and 30 HIV-negative) for the presence of BKV DNA. Samples from 3/15 (20 %) HIV-positive (one HIV-2 and two HIV-1) and 2/30 HIV-negative (6·6 %) individuals had positive results. This was comparable to the percentage of BKV-positive urine samples reported in HIV-positive (Sundsfjord *et al.*, 1994) and HIV-negative individuals (Tsai *et al.*, 1997). DNA extracted from fresh-frozen autopsy samples of the heart and striated muscle of the BKV vasculopathy patient, as well as DNA extracted from the urine of one HIV-2-positive patient and one HIV-negative healthy control subject who tested positive by qualitative BKV PCR, were subjected to long PCR amplification.

Long PCR amplification of full-length BKV genomes

BKV_{DUN} genomic DNA consists of a single copy of double-stranded circular DNA of 5153 bp (Seif *et al.*, 1979). We therefore adapted a long PCR technique initially developed for JCV (Agostini & Stoner, 1995) and SV40 (Lednicky *et al.*, 1997) to amplify full-length BKV genomes in a single reaction. Amplification of BKV DNA extracted from the heart and muscle of the patient with BKV vasculopathy and from the urine samples of one HIV-2-positive patient and one healthy control subject each yielded a 5 kb fragment. To study the potential genetic diversity of these viral isolates, amplified fragments were cloned and sequenced. Full-length sequences were obtained from five BKV clones from the muscle and four BKV clones from the heart of the patient with BKV vasculopathy (BKV_{CAP}) and from three clones from the urine of the HIV-2-positive patient (BKV_{HI}) and the healthy control subject (BKV_{HC}), respectively (Table 2).

Variation within the regulatory region

The regulatory regions of BKV_{CAP}, BKV_{HI} and BKV_{HC} consistently had an O_{1-142}-P_{1-68}-Q_{1-39}-R_{1-63}-S_{1-63} pattern identical to the archetype BKV_{WW} regulatory region (Markowitz & Dynan, 1988; Moens *et al.*, 1995; Sugimoto *et al.*, 1990) in all clones, except for minor deletions from nt 182 to 190 in one BKV_{HI} clone (HI-u8) and from nt 224 to 272 in one BKV_{HC} clone (HC-u2), which did not affect the T antigen binding sites (Moens *et al.*, 1995). Single isolated nucleotide substitutions were also found in CAP-m13 (T→C at nt 45) and CAP-m5 (A→G at nt 278), the latter being also present in HI-u6 and HI-u8 clones. This indicates that, unlike other BKV strains with expanded tropism (Stoner *et al.*, 2002), infection of capillary endothelial cells by BKV_{CAP} was not caused by changes in its

```
                     1         2         3         4         5         6
            123456789012345678901234567890123456789012345678901234567890123
                     |         |         |         |         |         |
CAP cons    TAAAATTAGAAAACTTTAATTAAAAAAATCAATTATTAATTTTAATAAATATTTTCAAATTTA
CAP-m2      .....C.....G....................C.......C......G....A..G.GC....    9
CAP-m5      .G..........................C..G.......C..........G............    5
CAP-m9      .......G.....TC..G.CC........T...C..CG.....G.......C..C.T......   14
CAP-m13     C...G..........C........G.G........C......C.....G....C.......C.   10
CAP-m18     ...G..C.AG......C......G...C......C.........GCG...G.....T...C..   14
CAP-h2      ..........G.......G...G.......................................G    4
CAP-h5      ..T.........G............G............G........................    4
CAP-h8      .....................C...................C.......C.............    3
CAP-h22     ..............................G............................C...    2
```

Fig. 1. Variable sites among the CAP clones. Each sequence is compared with the consensus, and only sites with differences from the consensus are shown. The total number of differences from the consensus is shown on the right. Within the alignment of 5141 nt, the positions of the 63 variable sites are: 45, 278, 402, 529, 680, 754, 1139, 1238, 1412, 1561, 1606, 1869, 1911, 1913, 1931, 1941, 2326, 2414, 2417, 2462, 2483, 2540, 2609, 2610, 2664, 2769, 2809, 2912, 2991, 3047, 3087, 3241, 3287, 3358, 3374, 3387, 3509, 3646, 3670, 3680, 3798, 3868, 3886, 3936, 3980, 4243, 4254, 4282, 4342, 4375, 4401, 4567, 4623, 4691, 4821, 4851, 4864, 4884, 4954, 5017, 5032, 5057.

regulatory region. In addition, six isolated substitutions were found in all HC clones compared with BKV_{WW} including T→G at nt 41, G→A at nt 86, A→T at nt 160, T→C at nt 173, A→G at nt 253 and A→G at nt 330.

Variation within the coding region

The coding region sequences of the BKV_{CAP}, BKV_{HI} and BKV_{HC} strains had unique nucleotide differences in each clone. There was an unusual pattern of sequence variation among the nine CAP clones (Fig. 1). At 61 of 63 variable sites, one clone differed from the other eight; at the other two sites, the difference was shared by just two clones. The lack of variation shared among multiple clones has implications concerning the structure of the viral population within a host (discussed below). It also implies that the consensus sequence is the best estimate of the sequence of the common ancestor of the various clones. Interestingly, the five muscle clones exhibited greater divergence from this consensus (mean 10·4 differences) than the four heart clones (mean 3·3). The differences from the consensus can be used to examine the pattern of mutation within the BKV population. These mutations were highly non-random (Table 3). The vast majority (59/63 = 94 %) of changes occurred at A or T sites. In addition, most of the changes (55/63 = 87 %), including most of the changes at A or T sites (52/59 = 88 %), were transitions. Of note, it was not possible to distinguish, for example, whether an A→G change on the strand sequenced represented an A→G mutation on that strand or a T→C mutation on the complementary strand. Thus, it was possible to distinguish six categories of mutation: among the CAP clones, 52 of the total of 63 changes (83 %) belonged to just one category, transitions from A or T.

Interestingly, the coding region sequences of the multiple CAP, HI and HC clones each contained unique nucleotide substitutions (Table 4). The minimum number of differences between any pair of sequences from the same patient was five. The maximal difference was 26, or 0·55 % per site, between two CAP clones. On average, about twice as much variation was seen among the CAP sequences (0·29 %) as among those from HI (0·16 %) or HC (0·14 %), but it was not clear whether this reflected greater diversity in that patient or was a consequence of sampling only from urine in the case of HI and HC. Within the BKV_{CAP} clones, the nucleotide changes were irregularly spread throughout the coding region and were maximum in the small t gene, followed by the large T, VP1, agnoprotein, VP2 and VP3 genes.

To determine whether errors of the *Taq* Plus Long PCR system could be responsible for the apparent genetic diversity of the BKV samples, the full-length BKV_{DUN} DNA was amplified using the same method in 10 individual PCRs. A 5153 bp PCR-amplified product was cloned into the pCR2.1-TOPO vector and a 1360 bp fragment (nt 2154–3513) from one clone of each PCR was sequenced. This fragment was chosen because it contains 20 point mutations spread among nine BKV_{CAP} clones ($1·63 \times 10^{-3}$

Table 3. Inferred nucleotide changes among CAP clones

Changes were inferred from the sequenced strand; complementary changes (see text) are shown on the same row.

Change	*n*	Change	*n*
A→G	27	T→C	25
A→C	4	T→G	0
A→T	2	T→A	1
C→T	2	G→A	1
C→G	1	G→C	0
C→A	0	G→T	0

Table 4. Nucleotide sequence divergence among BK virus clones

For each comparison, the upper values give the range of nucleotide differences, the lower value gives the mean percentage pairwise difference (per site). The total refers to the complete genome sequence, excluding the regulatory region between the T antigen gene and the agnoprotein gene. Numbers in parentheses denote the lengths of sequences (nt).

Strains	Gene						Total (4766)
	Agno (201)	VP2 (1056)	VP3 (699)	VP1 (1089)	LTAg (2088)	StAg 519)	
(a) Between clones from a single patient							
CAP (n=9)	0–2	0–5	0–3	0–8	2–13	0–5	5–26
	0·22	0·17	0·16	0·31	0·33	0·46	0·29
HI (n=3)	0	2–3	1–2	2–5	0–1	0–1	6–8
	0·00	0·25	0·19	0·37	0·03	0·13	0·16
HC (n=3)	0–1	1–3	1–3	0–1	1–4	0–1	5–8
	0·33	0·19	0·29	0·06	0·13	0·13	0·14
(b) Between clones from different patients							
CAP vs HI	0–1	3–7	2–5	7–13	5–11	0–4	17–29
	0·09	0·40	0·46	0·83	0·37	0·30	0·46
CAP vs HC	2–4	6–10	6–9	13–19	25–33	4–8	53–68
	1·25	0·68	0·98	1·38	1·37	0·98	1·26
HI vs HC	2–3	9–11	8–10	14–16	26–30	4–6	58–63
	1·13	0·92	1·29	1·35	1·33	0·90	1·27

differences per site). Comparison among the 10 BKV_{DUN} clones revealed only one nucleotide difference, found in a single clone (T→C at nt 3133 in PCR 7). Therefore, our estimate of the error rate of this long PCR system was 1/13 600 or $7{\cdot}35\times10^{-5}$ mutations per site for this region of the genome. In comparison with the published BKV_{DUN} sequence, all 10 of our clones differed by a single synonymous nucleotide substitution (A→T at nt 3472).

Phylogenetic analysis and genotypic classification of BKV sequences

The BKV strains from these three patients were compared with the three previously published full-length sequences in a phylogenetic analysis (Fig. 2a). Even though the extent of diversity among different CAP clones overlapped that between CAP and HI (Table 4), the multiple clones from each patient formed monophyletic clades, suggesting that each patient was infected from a single source.

Four serotypes of BKV have been described (Jin *et al.*, 1993b). In addition, four genotypes of BKV have been defined on the basis of nucleotide differences within a 23 codon 'typing' region of the VP1 gene that appears to encode the epitope responsible for serotypic differences. Most of these differences are non-synonymous, leading to amino acid variations at 11 sites. These genotypes appear to correlate with the four serotypes (Jin *et al.*, 1993a). The previously published full-length sequences represent genotypes I (DUN and MM) and III (AS). As expected, strain AS was highly divergent from strains MM and DUN in the evolutionary tree (Fig. 2a). The strains from the three patients described here were more closely related to each other than to any of the three previously described strains and much more closely related to genotype I than to genotype III. The CAP, HI and HC sequences all shared the same amino acids as genotype I at the sites previously used to define genotypes.

Full-length VP1 sequences were available from an additional six viruses. These included one sequence (GenBank accession no. Z19534) described as being the original BKV isolate (Jin *et al.*, 1993a). Preliminary comparisons indicated that this sequence was recombinant: across the first 320 codons this sequence differed from strain MM at only 2 nt and from strain SB at 43 nt, whereas in the other 43 codons it differed from MM at 11 nt but was identical to the SB sequence. Therefore, we excluded this sequence from the phylogenetic analysis of VP1 sequences (Fig. 2b). The other five additional sequences included one, SB, previously designated as genotype II (Gibson & Gardner, 1983); the other four appeared to be genotype I and consisted of WW (Chauhan *et al.*, 1984), DIK (Goudsmit *et al.*, 1981), JL (Pauw & Choufoer, 1978) and MT (Sugimoto *et al.*, 1989).

The VP1 phylogeny indicated that three of these genotype I sequences fell within the radiation of the strains CAP, HI and HC; surprisingly, one strain (JL) fell within the radiation of HC clones. JL differed from the HC-u5 and HC-u9 sequences at only one site, while HC-u2 differed from HC-u5, HC-u9 and JL at a single other site. This degree of identity among sequences from two individuals in different countries is surprising given the amount of diversity seen among the CAP clones (Tables 2 and 4).

The other interesting feature of the VP1 phylogeny was that the genotype II (SB) and III (AS) strains exhibited less

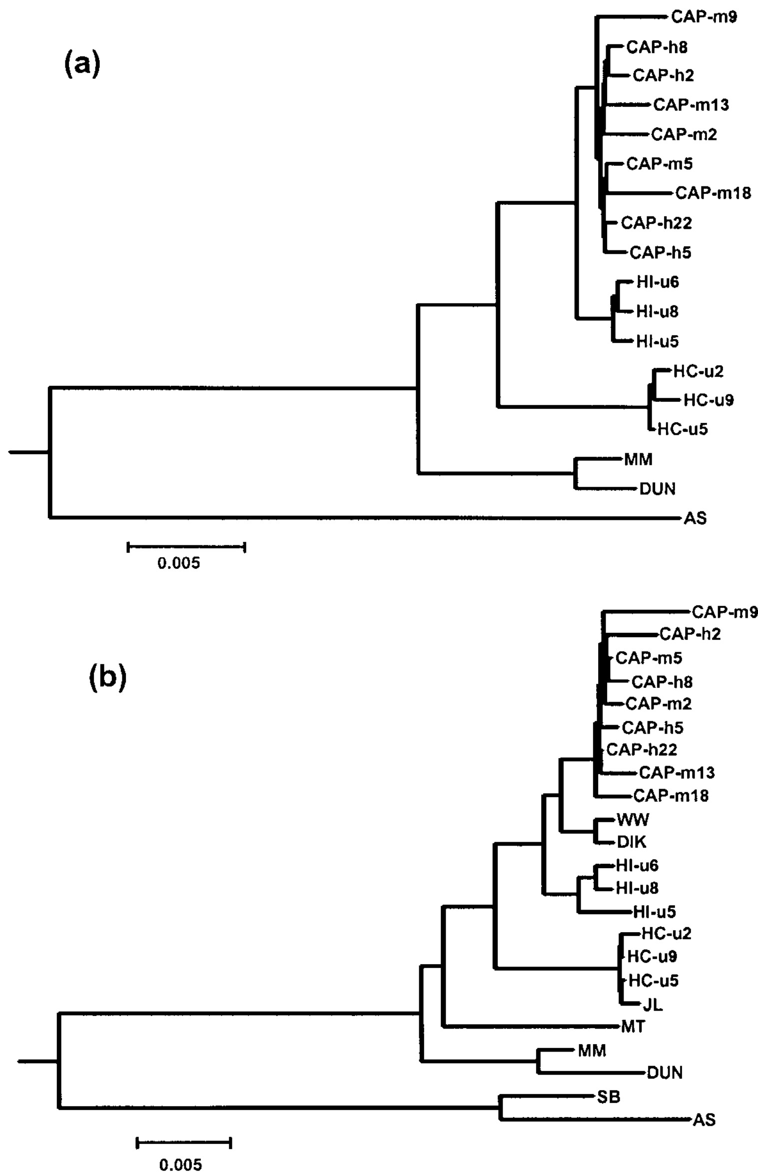

Fig. 2. Midpoint rooted maximum-likelihood phylogenetic trees of BKV sequences. (a) Tree derived from analysis of the complete nucleotide coding sequence of 18 BKV strains, excluding the non-coding regulatory region located between the T gene and the agnoprotein gene. (b) Tree derived from analysis of the VP1 gene sequences from 23 BKV strains. The scale bar indicates the scale of the branches in inferred numbers of substitution per site. Thus, the bars, of length 0·005, correspond to approximately 25 substitutions across the entire genome.

divergence than different genotype I strains. The various BKV strains included in the VP1 phylogeny were obtained from individuals in Africa, Asia, Europe and North America (Table 2). SB and AS were both derived from individuals in the UK, but otherwise among the genotype I strains there seemed to be no correlation between geographical origin and phylogenetic relationships.

Recently, it has been suggested that two subgroups of genotype I (Ia and Ib) can be defined on the basis of differences at three nucleotide positions within a larger region surrounding the typing region (Stoner *et al.*, 2002). However, all three sites are third positions of codons and in each case the differences proposed to be diagnostic of subgroups Ia and Ib were synonymous changes. Thus, these three differences cannot lead to any serological distinction. Furthermore, because synonymous substitutions generally occur more frequently than non-synonymous substitutions, these sites may not be reliable phylogenetic markers. Consideration of these sites within the VP1 sequences of strains in Fig. 2(b) confirmed this. The DUN and WW strains represent the proposed subgroups Ia and Ib: DUN had T-G-T at the three sites, compared with A-A-C in strain WW. Strain MT had A-G-T, different from both of the other two; since MT is phylogenetically distinct from DUN and WW (Fig. 2b), it might seem reasonable to classify this strain as a third subgroup. However, while eight of the nine CAP clones had A-A-C (as for subgroup Ib), the other (CAP-m9) had A-A-T; clearly this clone should not be classified as a fourth subgroup. The CAP-m9 clone fell within the radiation of CAP sequences, but at these three sites differed from other CAP sequences to the same degree that it differed from MT or even from genotypes II and III (both A-C-T). Thus, variations at these three sites do not seem to form a sound basis for classification into subgroups.

Analysis of VP1 proteins

BKV_{CAP} tropism for vascular endothelial cells may have been caused by changes in the amino acid sequence composition of the major capsid protein, VP1. We therefore aligned the deduced VP1 sequences of all BKV_{CAP} clones and compared them with BKV_{DUN}, as well as the novel BKV_{HI} and BKV_{HC} clones (Table 5a). BKV_{CAP} clones had three to seven VP1 amino acid changes compared with BKV_{DUN}. Some of these amino acid changes were unique to a given clone, but none of them was present in more than one of the BKV_{CAP} clones and not in the BKV_{HI} and BKV_{HC} clones.

On the contrary, three amino acid changes were present in all BKV_{HC} clones only (aa 42, V→L; aa 210, V→I; and aa 362, L→V). In addition, one BKV_{HI} clone (HI-u5) had a single mutation at nt 1702 (G→C), leading to an amino acid change from glutamic acid to glutamine at position 73 of the VP1 protein, resulting in the loss of a negatively charged amino acid side chain. This had been previously described in the partially sequenced BKV_{Yale} strain, which was isolated

Table 5. Sequence variations within the predicted amino acid sequences of (a) the VP1 protein and (b) the large T protein of BKV CAP, HI and HC clones compared with BKV_{DUN} strain (only differences are shown)

(a)

Site*	4	42	47	73	121	127	146	158	170	171	175	210	219	259	288	289	304	305	311	313	330	353	362
BKV_{DUN}	T	V	E	E	I	L	Q	E	R	S	D	V	A	S	K	I	F	L	N	R	E	K	L
CAP-m2	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	–	–
CAP-m5	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	–	–
CAP-m9	–	–	–	–	T	–	–	D	–	T	–	–	T	G	–	T	S	–	–	–	–	–	–
CAP-m13	–	–	–	–	–	P	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	–	–
CAP-m18	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	R	–
CAP-h2	A	–	–	–	–	–	–	D	–	T	–	–	T	–	R	–	–	–	–	–	G	–	–
CAP-h5	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	–	–
CAP-h8	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	T	–	–	–	–
CAP-h22	–	–	–	–	–	–	–	D	–	T	–	–	T	–	–	–	–	–	–	–	–	–	–
HI-u5	–	–	Q	–	–	–	R	D	–	T	E	–	T	–	–	–	–	–	–	–	–	–	–
HI-u6	–	–	–	Q	–	–	–	D	–	T	E	–	T	–	–	–	–	S	–	–	–	–	–
HI-u8	–	–	–	Q	–	–	–	D	W	T	E	–	T	–	–	–	–	–	–	–	–	–	–
HC-u2	–	L	–	–	–	–	–	D	–	T	E	I	T	–	–	–	–	–	–	G	–	–	V
HC-u5	–	L	–	–	–	–	–	D	–	T	E	I	T	–	–	–	–	–	–	–	–	–	V
HC-u9	–	L	–	–	–	–	–	D	–	T	E	I	T	–	–	–	–	–	–	–	–	–	V

(b)

Site*	20	28	37	42	63	133	142	182	260	267	273	288	323	327	334	353	373	430	471	475	504	584	603	629	675
BKV_{DUN}	E	P	K	H	E	K	S	S	S	T	K	F	F	I	N	T	K	I	M	E	D	F	E	D	Q
CAP-m2	–	–	–	–	–	–	–	P	N	–	–	–	–	T	–	–	R	–	–	–	–	–	–	–	–
CAP-m5	–	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	G	–
CAP-m9	–	–	–	–	–	E	–	–	N	–	–	–	–	T	–	–	–	T	V	–	G	–	K	–	–
CAP-m13	–	–	R	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	G	–	–	–	–	–
CAP-m18	–	–	–	R	–	–	–	–	N	–	R	L	–	T	–	–	–	–	–	–	–	–	–	–	–
CAP-h2	–	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	–
CAP-h5	–	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	–
CAP-h8	–	–	–	–	–	–	–	–	N	–	–	–	–	T	D	–	–	–	–	–	–	–	–	–	–
CAP-h22	–	–	–	–	G	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	S	–	–	–
HI-u5	–	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	E
HI-u6	G	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	E
HI-u8	–	–	–	–	–	–	–	–	N	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	E
HC-u2	–	S	–	–	–	–	C	–	N	–	–	–	–	T	–	S	–	–	–	–	–	–	–	–	–
HC-u5	–	S	–	–	–	–	C	–	N	–	–	–	–	T	–	S	–	–	–	–	–	–	–	–	–
HC-u9	–	S	–	–	–	–	C	–	N	A	–	–	C	T	–	S	–	–	–	–	–	–	–	–	–

*m, Striated muscle; h, heart; u, urine.

from a leukaemia patient with BKV tubulointerstitial nephritis and meningoencephalitis. This amino acid change could possibly promote a β-strand in this region at the expense of the α-helix (Stoner *et al.*, 2002).

The only crystal structure of the VP1 capsid protein of polyomaviruses that has been characterized is that of SV40 (Liddington *et al.*, 1991). We therefore examined the structure of SV40 VP1 pentamers at positions that correspond to the amino acid changes seen in the VP1 protein of BKV_{CAP} clones compared with BKV_{HI} and BKV_{HC}. None of these changes were at positions likely to be exposed on the surface of the virions and, therefore, no conformational differences could be expected to result in a novel interaction with potential cellular receptors. In addition, the C terminus of the capsid protein VP2, which contains its binding site with the VP1 pentamers, showed no significant differences among the clones of these three viruses.

Analysis of T antigen sequences

The large T protein is the main regulatory protein of BKV. It has a crucial role in the life cycle of this virus since its presence is required for the initiation of BKV DNA replication and for activation of the switch from early to late gene transcription. It also contains binding sites for nuclear proteins, as well as a host-range domain. Therefore, changes in the T antigen structure may be responsible for the vascular tropism of the BKV_{CAP} strain.

Although three to seven amino acid changes compared with BKV_{DUN} could be seen in the T antigen of BKV_{CAP} clones, none of them was present in all BKV_{CAP} clones, nor in the BKV_{HI} or BKV_{HC} clones (Table 5b). Analysis of the T antigen binding site showed no mutations at positions 4298 and 4299, previously described in the virulent BKV_{Cin} strain in an AIDS patient with end-stage renal disease (Smith *et al.*, 1998). In addition, no significant amino acid changes were detected in any of the BKV_{CAP} clones in the DNA-binding domain, zinc-finger domain, p53-binding domains, host range domains or phosphorylation sites. Therefore, changes of the amino acid sequence of the T antigen of BKV_{CAP} clones could not readily explain the difference in cell tropism of this novel BKV strain. In contrast, three unique amino acid changes were found in the BKV_{HC} clones only and one amino acid change was found in BKV_{HI} clones only.

Analyses of amino acid sequences of the t, VP2, VP3 and agnoproteins

The deduced amino acid sequences of the t, VP2, VP3 and agnoproteins of all the BKV_{CAP}, BKV_{HI} and BKV_{HC} clones were aligned and compared with the BKV_{DUN} amino acid sequence. Unique changes were found only in BKV_{HC} clones, in the small t protein (aa 28, P→S), the VP2 (aa 318, Q→K), VP3 (aa 199, Q→K) and the agnoprotein (aa 14, V→L).

DISCUSSION

Only three full-length BKV sequences have been published to date. These, as well as most partially sequenced BKV strains, have been characterized by co-cultivation of the patient's primary isolates in permissive cells, followed by molecular cloning and sequencing of a single clone. This could promote the selection of tissue culture-adapted strains bearing artificial molecular changes, especially at the level of their regulatory region. Furthermore, potential intra-strain polymorphisms cannot be evaluated.

To circumvent the need for virus culture altogether, we chose to amplify the complete genome of BKV isolates in a single PCR. Moreover, we sequenced three to nine clones from each patient's isolates, to gain insights regarding their genetic diversity. The availability of 15 novel full-length molecular clones of BKV thus radically expands our knowledge of the genetic diversity of this virus. In addition, comparison of multiples clones derived simultaneously from the same patient sheds a new light on BKV evolution. Indeed, there are few data available on the intra-patient genetic diversity of the coding region of BKV, since previously published studies used direct sequencing of PCR products, which aims to reveal only the most prominent species present in a sample and renders the detection of polymorphisms difficult (Jin *et al.*, 1993b; Randhawa *et al.*, 2002).

The extent of sequence variation among clones from individual patients in our study was surprising. The most divergent sequences from patient CAP differed at 0·55 % of sites. BKV infections often occur at a very early age (Tavis *et al.*, 1990). This patient with BKV vasculopathy was 52 years old when the samples were taken. If the initial BKV infection arose from a single strain, two viruses from this BKV population could have been diverging for about 50 years. A divergence of 0·55 % then implies a rate of nucleotide substitution of around 5×10^{-5} substitutions per site per year [i.e. 0·55 %/(2×50) to account for divergence along two lineages from the common ancestor]. Even taking the lower mean divergence seen in the two other patients (i.e. around 0·15 %) who could each have been infected for a little over 40 years yields a rate of evolution around 2×10^{-5} substitutions per site per year. These rate estimates are surprisingly high for a DNA virus and about two orders of magnitude greater than the rate of 4×10^{-7} per site per year estimated for synonymous substitutions among different JCV genotypes (Hatwell & Sharp, 2000).

Several factors could contribute to these results. First, the possibility that the genetic diversity observed among the BKV clones was artificially inflated by errors of the *Taq* polymerase during the PCR and by sequencing artefacts must be considered. However, the fidelity of the *Taq* Plus DNA polymerase we used for the long PCR amplification is extremely high, with an error rate of only $1{\cdot}3 \times 10^{-6} \pm 2{\cdot}2$ (Cline *et al.*, 1996). Therefore, on average, the PCR enzyme should give rise to only 1 nt change per 150 full-length BKV

clones under ideal conditions. However, this error rate may only pertain to experiments conducted under certain optimal conditions. In our hands, sequencing of long PCR-amplified products of the BKV_{DUN} reference strain from 10 individual reactions revealed a somewhat higher error rate ($7{\cdot}35\times10^{-5}$). Nevertheless, the rate of mutations observed in the same region among the nine BKV_{CAP} clones was more than 20 times higher than this ($1{\cdot}63\times10^{-3}$) and thus this high intra-strain genetic diversity cannot be accounted for by errors of the *Taq* polymerase. Furthermore, sequencing artefacts could be discounted because the analyses were all performed on both strands of each clone. Surprisingly, we also found a single point mutation compared with the published BKV_{DUN} sequence (position 3472, silent mutation A→T) in all 10 clones. We believe that this mutation, which is present in all BKV_{DUN} clones from 10 separate PCRs, was not caused by an error of the PCR enzyme, but existed already in the plasmid obtained from the ATCC. Indeed, this A→T change has already been reported in BKV strains MM and AS, and was seen in this study in all CAP, HI and HC clones.

Secondly, although there are no obvious aspects of BKV and JCV biology that would be expected to lead to large differences in their evolutionary rates, the rate of evolution of BKV may be higher than that of JCV because of a higher mutation rate and/or a higher replication rate. Consistent with this, the diversity among BKV strains is higher than that among JCV strains; for BKV, the differences among genotypes are extensive enough to give rise to distinct serotypes, whereas that is not the case for JCV. The extent of divergence between BKV genotypes I and III is about twice that seen between the most divergent JCV genotypes, which is only a little higher than that seen within BKV genotype I (Table 6). In contrast to JCV, there is as yet no clear association between genetic diversity and geographical origin or ethnic background for BKV strains, perhaps because the broader host cell range of BKV renders it more readily transmissible, and so the historical epidemiology of the two viruses may have been different. Therefore, the greater diversity in BKV than in JCV may merely reflect a greater length of time since the most recent common ancestor for BKV strains, rather than a faster evolutionary rate.

A third possibility is that the rates of BKV and JCV evolution are broadly similar, but that estimates of evolutionary rates are higher over short-term comparisons (i.e. within an individual for BKV) than across longer-term comparisons such as those between different JCV genotypes that may have diverged 100 000 years ago. This could arise due to saturation effects, such that longer-term comparisons give rise to underestimates, as previously discussed for certain viruses with RNA genomes (Sharp *et al.*, 2000; Simmonds, 2001). The observation that the transition-to-transversion ratio was higher among BKV/CAP clones than the value estimated for the phylogenetic analysis of all BKV sequences would be consistent with some saturation of transitions. Clearly, it will be of interest to compare the intra-host diversity seen for BKV with that for JCV.

Our analyses also revealed that the previously published VP1 gene sequence described as being from the original BKV isolate (Jin *et al.*, 1993a) is a likely recombinant of genotype I and genotype II sequences. Dual infection with more than one subtype has been described (Jin *et al.*, 1995), which may permit recombination to occur. A recombinant JCV sequence has been identified (Hatwell & Sharp, 2000). However, for the BKV sequence, as with the JCV sequence, it is not yet clear whether the sequence reflects a recombinant virus or 'recombination' in the laboratory.

Extensive analyses of potential molecular determinants of the vascular tropism of BKV_{CAP} remained inconclusive.

Table 6. Diversity among polyomavirus strains

L is the length of the BKV gene in codons. Ka and Ks are the estimated numbers of non-synonymous and synonymous substitutions per site (by the method of Li, 1993). BKV-1 values refer to the mean of comparisons between the CAP clones and the Dunlop strain. BKV-2 values refer to the mean of comparisons between Dunlop, CAP, HI and HC clones and AS. JCV values are for genotype 1 vs genotype 6 (taken from Hatwell & Sharp, 2000). The mean (weighted by gene length) is calculated from the non-overlapping genes for the agnoprotein, VP2, VP1 and large T antigen.

	L	BKV-1		BKV-2		JCV	
		Ka	Ks	Ka	Ks	Ka	Ks
Agnoprotein	66	0·010	0·035	0·014	0·076	0·015	0·000
VP2	351	0·011	0·026	0·015	0·113	0·008	0·048
VP3	232	0·007	0·032	0·018	0·144	0·012	0·066
VP1	362	0·005	0·066	0·027	0·176	0·008	0·110
Large T antigen	695	0·003	0·075	0·009	0·179	0·005	0·089
Small t antigen	172	0·004	0·039	0·008	0·175	0·003	0·043
Mean		0·006	0·059	0·015	0·158	0·007	0·080

Our results indicate that the regulatory region is not responsible for the new phenotype of this isolate. Indeed, the regulatory region of all nine BKV_{CAP} clones sequenced had an O-P-Q-R-S pattern identical to the archetype BKV_{WW} regulatory region (Markowitz & Dynan, 1988; Moens *et al.*, 1995; Sugimoto *et al.*, 1990) found in primary urine isolates, including the novel BKV_{HI} and BKV_{HC} strains described in this study. This contrasts with JCV, where central nervous system strains usually have a regulatory region with a typical tandem repeat pattern and variable deletions compared with the archetype regulatory region found in the kidney (Major *et al.*, 1992).

We also explored whether amino acid changes in the capsid proteins of the BKV_{CAP} clones could be responsible for conformational differences that may lead to their preferential entry in capillary endothelial cells. Such dramatic extension of host cell range has been demonstrated in mouse polyomavirus, where a single amino acid change (G→E) at position 92 in the VP1 protein induces an increase in tumour incidence in a broader spectrum of tissues (Freund *et al.*, 1991). However, we could not identify amino acid changes present in the VP1 or VP2 proteins of most BKV_{CAP} clones that were not previously identified in non-vasculotropic BKV strains (Table 5). In addition, conformational analysis of the VP1 protein of BKV_{CAP} clones failed to show any changes at the surface of the virions that could result in a novel interaction with potential cellular receptors. To our surprise, we found that each BKV_{CAP} clone had several unique amino acid differences. Since BKV_{CAP} was not detected in renal tubular cells but exclusively in the capillary endothelial cells of the kidney, muscle and heart of this patient, it is unclear which, if any, of these changes is responsible for the initial vascular tropism of this particular isolate. It is indeed possible that some of these changes are the consequence, instead of the cause, of the expanded host cell range of this virus.

In addition to cell entry, a productive infection by polyomaviruses requires a successful interaction between the cellular machinery and both the virus regulatory region and the T antigen, which is necessary to initiate the expression of capsid proteins and viral DNA replication. Again, no unique amino acid changes could be identified in the T antigen of all BKV_{CAP} clones that were not already described in non-vasculotropic isolates (Table 5b) and a similar diversity was found in their T antigen as in their VP1 protein (Table 5a).

Therefore, the event that led to successful entry and replication of BKV in capillary endothelial cells remains unclear. It may be that a unique feature of the patient himself, such as mutation in a cellular protein present in capillary endothelial cells, was in fact responsible for this peculiar clinical presentation. Furthermore, the mechanisms by which BKV caused an increased vascular permeability in this patient are not resolved. Direct action of the virus on the cell membrane or possibly an immune-mediated mechanism may play a role in this syndrome. Studies of the phenotype of BKV_{CAP} clones *in vitro* are currently in progress in our laboratory to address these issues.

ACKNOWLEDGEMENTS

We are thankful to Dr Stephen C. Harrison for his help in the conformational analyses of the VP1 capsid protein. This work was supported in part by Public Health Service grant NS 01919 and NS/AI 41198 to I. J. K.

REFERENCES

Agostini, H. T. & Stoner, G. L. (1995). Amplification of the complete polyomavirus JC genome from brain, cerebrospinal fluid and urine using pre-PCR restriction enzyme digestion. *J Neurovirol* **1**, 316–320.

Agostini, H. T., Ryschkewitsch, C. F., Singer, E. J. & Stoner, G. L. (1996). Co-infection with two JC virus genotypes in brain, cerebrospinal fluid or urinary tract detected by direct cycle sequencing of PCR products. *J Neurovirol* **2**, 259–267.

Agostini, H. T., Yanagihara, R., Davis, V., Ryschkewitsch, C. F. & Stoner, G. L. (1997). Asian genotypes of JC virus in Native Americans and in a Pacific Island population: markers of viral evolution and human migration. *Proc Natl Acad Sci U S A* **94**, 14542–14546.

Berger, J. R., Kaszovitz, B., Post, M. J. & Dickinson, G. (1987). Progressive multifocal leukoencephalopathy associated with human immunodeficiency virus infection. A review of the literature with a report of sixteen cases. *Ann Intern Med* **107**, 78–87.

Chauhan, S., Lecatsas, G. & Harley, E. H. (1984). Genome analysis of BK (WW) viral DNA cloned directly from human urine. *Intervirology* **22**, 170–176.

Cline, J., Braman, J. C. & Hogrefe, H. H. (1996). PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res* **24**, 3546–3551.

Coleman, D. V., Mackenzie, E. F., Gardner, S. D., Poulding, J. M., Amer, B. & Russell, W. J. (1978). Human polyomavirus (BK) infection and ureteric stenosis in renal allograft recipients. *J Clin Pathol* **31**, 338–347.

Coleman, D. V., Wolfendale, M. R., Daniel, R. A., Dhanjal, N. K., Gardner, S. D., Gibson, P. E. & Field, A. M. (1980). A prospective study of human polyomavirus infection in pregnancy. *J Infect Dis* **142**, 1–8.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.5c. Department of Genetics, University of Washington, Seattle, USA.

Ferrante, P., Mediati, M., Caldarelli-Stefano, R., Losciale, L., Mancuso, R., Cagni, A. E. & Maserati, R. (2001). Increased frequency of JC virus type 2 and of dual infection with JC virus type 1 and 2 in Italian progressive multifocal leukoencephalopathy patients. *J Neurovirol* **7**, 35–42.

Freund, R., Calderone, A., Dawe, C. J. & Benjamin, T. L. (1991). Polyomavirus tumor induction in mice: effects of polymorphisms of VP1 and large T antigen. *J Virol* **65**, 335–341.

Gardner, S. D., Field, A. M., Coleman, D. V. & Hulme, B. (1971). New human papovavirus (BK) isolated from urine after renal transplantation. *Lancet* **1**, 1253–1257.

Gibson, P. E. & Gardner, S. D. (1983). Strain differences and some serological observations on several isolates of human polyomaviruses. *Prog Clin Biol Res* **105**, 119–132.

Goudsmit, J., Baak, M. L., Sleterus, K. W. & Van der Noordaa, J. (1981). Human papovavirus isolated from urine of a child with acute tonsillitis. *Br Med J (Clin Res Ed)* **283**, 1363–1364.

Guo, J., Sugimoto, C., Kitamura, T. & 10 other authors (1998). Four geographically distinct genotypes of JC virus are prevalent in China and Mongolia: implications for the racial composition of modern China. *J Gen Virol* **79**, 2499–2505.

Hatwell, J. N. & Sharp, P. M. (2000). Evolution of human polyomavirus JC. *J Gen Virol* **81**, 1191–1200.

Hirsch, H. H., Knowles, W., Dickenmann, M., Passweg, J., Klimkait, T., Mihatsch, M. J. & Steiger, J. (2002). Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients. *N Engl J Med* **347**, 488–496.

Jin, L., Gibson, P. E., Knowles, W. A. & Clewley, J. P. (1993a). BK virus antigenic variants: sequence analysis within the capsid VP1 epitope. *J Med Virol* **39**, 50–56.

Jin, L., Gibson, P. E., Booth, J. C. & Clewley, J. P. (1993b). Genomic typing of BK virus in clinical specimens by direct sequencing of polymerase chain reaction products. *J Med Virol* **41**, 11–17.

Jin, L., Pietropaolo, V., Booth, J. C., Ward, K. H. & Brown, D. W. (1995). Prevalence and distribution of BK virus subtypes in healthy people and immunocompromised patients detected by PCR-restriction enzyme analysis. *Clin Diagn Virol* **3**, 285–295.

Johnsen, J. I., Seternes, O. M., Johansen, T., Moens, U., Mantyjarvi, R. & Traavik, T. (1995). Subpopulations of non-coding control region variants within a cell culture-passaged stock of BK virus: sequence comparisons and biological characteristics. *J Gen Virol* **76**, 1571–1581.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110–119.

Knowles, W. A., Gibson, P. E. & Gardner, S. D. (1989). Serological typing scheme for BK-like isolates of human polyomavirus. *J Med Virol* **28**, 118–123.

Koralnik, I. J., Boden, D., Mai, V. X., Lord, C. I. & Letvin, N. L. (1999). JC virus DNA load in patients with and without progressive multifocal leukoencephalopathy. *Neurology* **52**, 253–260.

Lednicky, J. A., Jafar, S., Wong, C. & Butel, J. S. (1997). High-fidelity PCR amplification of infectious copies of the complete simian virus 40 genome from plasmids and virus-infected cell lysates. *Gene* **184**, 189–195.

Li, W. H. (1993). Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* **36**, 96–99.

Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L. & Harrison, S. C. (1991). Structure of simian virus 40 at 3·8-A resolution. *Nature* **354**, 278–284.

Major, E. O., Amemiya, K., Tornatore, C. S., Houff, S. A. & Berger, J. R. (1992). Pathogenesis and molecular biology of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. *Clin Microbiol Rev* **5**, 49–73.

Markowitz, R. B. & Dynan, W. S. (1988). Binding of cellular proteins to the regulatory region of BK virus DNA. *J Virol* **62**, 3388–3398.

Martin, J. D. & Foster, G. C. (1984). Multiple JC virus genomes from one patient. *J Gen Virol* **65**, 1405–1411.

Moens, U., Johansen, T., Johnsen, J. I., Seternes, O. M. & Traavik, T. (1995). Noncoding control region of naturally occurring BK virus variants: sequence comparison and functional analysis. *Virus Genes* **10**, 261–275.

Nickeleit, V., Hirsch, H. H., Binet, I. F., Gudat, F., Prince, O., Dalquen, P., Thiel, G. & Mihatsch, M. J. (1999). Polyomavirus infection of renal allograft recipients: from latent infection to manifest disease. *J Am Soc Nephrol* **10**, 1080–1089.

Padgett, B. L., Walker, D. L., ZuRhein, G. M., Eckroade, R. J. & Dessel, B. H. (1971). Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet* **1**, 1257–1260.

Pauw, W. & Choufoer, J. (1978). Isolation of a variant of BK virus with altered restriction endonuclease pattern. *Arch Virol* **57**, 35–42.

Petrogiannis-Haliotis, T., Sakoulas, G., Kirby, J. & 9 other authors (2001). BK-related polyomavirus vasculopathy in a renal-transplant recipient. *N Engl J Med* **345**, 1250–1255.

Randhawa, P. S., Khaleel-Ur-Rehman, K., Swalsky, P. A., Vats, A., Scantlebury, V., Shapiro, R. & Finkelstein, S. (2002). DNA sequencing of viral capsid protein VP-1 region in patients with BK virus interstitial nephritis. *Transplantation* **73**, 1090–1094.

Rice, S. J., Bishop, J. A., Apperley, J. & Gardner, S. D. (1985). BK virus as cause of haemorrhagic cystitis after bone marrow transplantation. *Lancet* **2**, 844–845.

Seif, I., Khoury, G. & Dhar, R. (1979). The genome of human papovavirus BKV. *Cell* **18**, 963–977.

Shadan, F. F. & Villarreal, L. P. (1993). Coevolution of persistently infecting small DNA viruses and their hosts linked to host-interactive regulatory domains. *Proc Natl Acad Sci U S A* **90**, 4117–4121.

Sharp, P. M., Bailes, E., Gao, F., Beer, B. E., Hirsch, V. M. & Hahn, B. H. (2000). Origins and evolution of AIDS viruses: estimating the time-scale. *Biochem Soc Trans* **28**, 275–282.

Simmonds, P. (2001). The origin and evolution of hepatitis viruses in humans. *J Gen Virol* **82**, 693–712.

Smith, R. D., Galla, J. H., Skahan, K., Anderson, P., Linnemann, C. C., Jr, Ault, G. S., Ryschkewitsch, C. F. & Stoner, G. L. (1998). Tubulointerstitial nephritis due to a mutant polyomavirus BK virus strain, BKV(Cin), causing end-stage renal disease. *J Clin Microbiol* **36**, 1660–1665.

Soeda, E., Maruyama, T., Arrand, J. R. & Griffin, B. E. (1980). Host-dependent evolution of three papova viruses. *Nature* **285**, 165–167.

Stoner, G. L., Alappan, R., Jobes, D. V., Ryschkewitsch, C. F. & Landry, M. L. (2002). BK virus regulatory region rearrangements in brain and cerebrospinal fluid from a leukemia patient with tubulointerstitial nephritis and meningoencephalitis. *Am J Kidney Dis* **39**, 1102–1112.

Sugimoto, C., Hara, K., Taguchi, F. & Yogo, Y. (1989). Growth efficiency of naturally occurring BK virus variants in vivo and in vitro. *J Virol* **63**, 3195–3199.

Sugimoto, C., Hara, K., Taguchi, F. & Yogo, Y. (1990). Regulatory DNA sequence conserved in the course of BK virus evolution. *J Mol Evol* **31**, 485–492.

Sundsfjord, A., Flaegstad, T., Flo, R., Spein, A. R., Pedersen, M., Permin, H., Julsrud, J. & Traavik, T. (1994). BK and JC viruses in human immunodeficiency virus type 1-infected persons: prevalence, excretion, viremia, and viral regulatory regions. *J Infect Dis* **169**, 485–490.

Takemoto, K. K., Rabson, A. S., Mullarkey, M. F., Blaese, R. M., Garon, C. F. & Nelson, D. (1974). Isolation of papovavirus from brain tumor and urine of a patient with Wiskott–Aldrich syndrome. *J Natl Cancer Inst* **53**, 1205–1207.

Tavis, J. E., Walker, D. L., Gardner, S. D. & Frisque, R. J. (1989). Nucleotide sequence of the human polyomavirus AS virus, an antigenic variant of BK virus. *J Virol* **63**, 901–911.

Tavis, J. E., Frisque, R. J., Walker, D. L. & White, F. A., III (1990). Antigenic and transforming properties of the DB strain of the human polyomavirus BK virus. *Virology* **178**, 568–572.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment

through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Tsai, R. T., Wang, M., Ou, W. C. & 7 other authors (1997). Incidence of JC viruria is higher than that of BK viruria in Taiwan. *J Med Virol* **52**, 253–257.

Weber, T., Trebst, C., Frye, S. & 8 other authors (1997). Analysis of the systemic and intrathecal humoral immune response in progressive multifocal leukoencephalopathy. *J Infect Dis* **176**, 250–254.

Yang, R. C. & Wu, R. (1979). BK virus DNA: complete nucleotide sequence of a human tumor virus. *Science* **206**, 456–462.

Yasunaga, T. & Miyata, T. (1982). Evolutionary changes of nucleotide sequences of papova viruses BKV and SV40: they are possibly hybrids. *J Mol Evol* **19**, 72–79.