Conserved archetypal configuration of the transcriptional control region during the course of BK polyomavirus evolution

Yoshiaki Yogo,1 Shan Zhong,1 Yawei Xu,2 Mengyun Zhu,2 Yuegen Chao,3 Chie Sugimoto,4,5 Hiroshi Ikegaya,6 Ayako Shibuya1 and Tadaichi Kitamura1

1Department of Urology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
2Department of Cardiovascular Medicine, Shanghai Tenth People’s Hospital, School of Medicine, Tongji University, 301 Yan Chang Road, Shanghai 200072, PR China
3Department of Urology, Shanghai Tenth People’s Hospital, School of Medicine, Tongji University, 301 Yan Chang Road, Shanghai 200072, PR China
4AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan
5Tsukuba Primate Research Center, National Institute of Biomedical Innovation, 1 Hachimandai, Tsukuba, Ibaraki 305-0843, Japan
6Department of Forensic Medicine, Graduate School of Medicine, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

INTRODUCTION

BK polyomavirus (BKV) is widespread among humans, asymptptomatically infecting children and then persisting in renal tissue. The transcriptional control region (TCR) of the BKV genome is variable among clinical isolates. Thus, archetypal TCRs with a common basic configuration generally occur in BKV isolates from the urine of immunocompromised patients, but rearranged TCRs that possibly arise from the archetypal configuration have also been detected in clinical specimens. To examine the hypothesis that archetypal strains represent wild-type strains circulating in the human population (the archetype hypothesis), we analysed 145 complete viral genomes amplified directly from the urine of non-immunocompromised individuals worldwide. These genomes included 82, three, two and 58 sequences classified as belonging to subtypes I, II, III and IV, respectively. Rearranged TCRs with long duplications or deletions were detected from two subtype I and two subtype IV genomes, but not from the other 141 genomes (thus, the TCRs of these genomes were judged to be archetypal). The variations in the archetypal TCRs were nucleotide substitutions and single-nucleotide deletions, most of which were unique to particular subtypes or subgroups. We confirmed that the four complete BKV genomes with rearranged TCRs did not form a unique lineage on a phylogenetic tree. Collectively, the findings demonstrate that the archetypal TCR configuration has been conserved during the evolution of BKV, providing support for the archetype hypothesis. Additionally, we suggest that ‘archetype’ should be used as a conceptual term that denotes a prototypical structure that can generate various rearranged TCRs during viral growth in vivo and in vitro.
renal dysfunction such as BKV-associated nephropathy (Hirsch & Steiger, 2003).

The genome of BKV is a single molecule of double-stranded, covalently circular DNA of about 5100 bp in length, and consists of the early, late and regulatory regions (Seif et al., 1979). The regulatory region contains the origin of replication and the transcriptional control region (TCR). Before 1986, the structure of the BKV TCR was determined (Wu, 1979; Pater et al., 1983; Watanabe & Yoshiike, 1985; Pagnani et al., 1986). However, more recent studies (ter Schegget et al., 1985; Rubinstein et al., 1987; Sugimoto et al., 1989; Tavis et al., 1989) have identified BKV genomes that lack tandem repeats and contain an extra 63 bp sequence (also referred to as an R block, see below) not present in the first BKV isolate (Gardner et al., 1971). Two of these strains were directly isolated from the urine of a patient by using molecular cloning (Rubinstein et al., 1987; Sugimoto et al., 1989). Strains with tandem repeats grow efficiently in tissue culture, whereas those without tandem repeats grow poorly (Hara et al., 1986; Watanabe & Yoshiike, 1985).

Yoshiike & Takemoto (1986) and Rubinstein et al. (1987) proposed a hypothesis (herein designated the archetype hypothesis) that BKV strains (archetypal strains) lacking tandem repeats represent wild-type BKV circulating in the human population. However, direct detection methods such as PCR or direct molecular cloning have shown that both archetypal and rearranged BKV TCRs are present in tissue from various clinical sources (reviewed by Moens & Van Ghelue, 2005). These findings suggest that either BKV strains with rearranged TCRs are occasionally generated during infection of humans or both BKV strains with archetypal and rearranged TCRs are circulating in the human population. In addition, there are further problems with the archetype hypothesis: (i) archetypal TCRs of BKV have been demonstrated primarily from the urine of immunocompromised patients, and it is unclear whether such TCRs are widespread in healthy or immunocompetent individuals; (ii) generally, archetypal BKV TCRs in subgenomic fragments have been PCR-amplified and therefore it is possible that they were derived from partial or incomplete genomes; and (iii) there is no direct evidence that an archetypal TCR can change to a rearranged TCR in vivo as well as in vitro (Imperiale & Major, 2007).

To examine the archetype hypothesis, here, we characterized the TCRs of 145 complete genomes of BKV (all of which were PCR-amplified) obtained in this and previous studies from urine of non-immunocompromised individuals worldwide. The 145 genomes included 49 new sequences isolated in this study and 96 previously reported sequences (Nishimoto et al., 2007; Yogo et al., 2007; Zheng et al., 2007; Zhong et al., 2007a), and comprised 82 of subtype I (prevalent throughout the world), 58 of subtype IV (prevalent in Asia and Europe), two and three of subtypes II and III (both of which are rare worldwide), respectively. Since the divergence of BKV into various subtypes occurred a long time ago (probably earlier than the emergence of modern humans) (Nishimoto et al., 2006), we considered that a comparison of TCR structures among different subtypes would allow us to elucidate conservation of the architecture of the BKV TCR during evolution, as a test of the archetype hypothesis.

METHODS

Urine samples. Urine samples were collected from donors from different geographical regions of Japan (Fukuoka, Hirokasi, Kazuno, Koromogawa, Nagareyama, Nara and Okinawa), Saudi Arabia (Riyadh) and China (Shanghai and Changchun). The Shanghai specimens were newly collected in this study at the Shanghai Tenth People’s Hospital, School of Medicine, Tongji University, but those from Japan, Saudi Arabia and Changchun were collected during previous studies (Sugimoto et al., 1997; Kitamura et al., 1998; Zheng et al., 2004). The donors of the urine samples were native inhabitants of each region (healthy volunteers or general patients without immunosuppression) aged 40 years or older (from Riyadh, Saudi Arabia) or 60 years or older (from Japan and Shanghai, China). All urine samples were collected with informed consent.

DNA extraction. The urine samples were centrifuged at 1300 g for 10 min at 4 °C to generate a pellet containing cell-associated viruses (PPT-1) and supernatant (SUP-1) containing cell-free viruses. PPT-1 was resuspended in 1.3 ml 10 mM Tris/HCl and 10 mM EDTA (pH 7.6) and digested with 100 μg proteinase K (Takara) ml⁻¹ in the presence of 0.5% SDS at 56 °C for 1 h. DNA was extracted from the digest and recovered by ethanol precipitation (Kitamura et al., 1990). The resultant pellet was dissolved in 80 μl Ultraspec Water (Biotec Laboratories). SUP-1 was centrifuged at 25 000 r.p.m. at 4 °C for 3 h in a Beckman SW28 swing rotor, and the resultant pellet (PPT-2) was processed as described for PPT-1. PPT-1 was used to extract DNA from urine samples of donors from Shanghai, China, and PPT-2 was used to extract DNA from urine samples of donors from Japan, Saudi Arabia and Changchun, China.

PCR. The entire BKV genome was amplified using Phusion high-fidelity DNA polymerase (Finnzymes) with the primer sets shown in Supplementary Table S1 (available in JGV Online). Amplified whole genomes were purified with a DNA sequencing. Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Amplified whole genomes were purified with a DNA extraction.

Phylogenetic analysis. The origins of 145 BKV genomes, which included 49 new sequences isolated in this study and 96 previously reported sequences (Nishimoto et al., 2007; Yogo et al., 2007; Zheng et al., 2007; Zhong et al., 2007a), are shown in Supplementary Table S2 (available in JGV Online). Whole BKV genome sequences excluding TCR sequences were aligned using CLUSTAL W (Thompson et al., 1994) with manual correction. Phylogenetic relationships among DNA sequences were evaluated using the neighbour-joining (NJ) method (Saitou & Nei, 1987). NJ analysis was performed under
Kimura’s two-parameter distance method (Kimura, 1980) using the baboon polyomavirus SA12 sequence (GenBank accession no. AY614708) as the outgroup (Cantalupo et al., 2005), and the phylogenetic tree was visualized using the NJ plot program (Perrière & Guoy, 1996). To assess the confidence level of the phylogenetic tree, bootstrap probabilities (BPs) were estimated with 1000 bootstrap replicates (Felsenstein, 1985).

**Statistical analysis.** Statistical analysis was performed using a Fisher’s exact probability test. The significance level was set at 5%.

**RESULTS**

**Classification of BKV genomes based on complete coding sequences**

We amplified and sequenced 49 complete BKV genomes from the urine of non-immunocompromised individuals in China (n = 24), Japan (n = 24) and Saudi Arabia (n = 1) (Supplementary Table S2). Combination of these sequences and complete BKV genome sequences (n = 96) previously determined from urine samples of non-immunocompromised individuals worldwide (Supplementary Table S2) yielded 145 complete BKV genome sequences. The entire coding sequences derived from the 145 BKV genomes, together with eight reference coding sequences [S, DUN, JL, KOM-7, MT, RYU-3, TW-3 and WW (Seif et al., 1979; Tavis et al., 1989; Nishimoto et al., 2006)], were subjected to reconstruction as an NJ phylogenetic tree using SA12 (a primate polyomavirus-related BKV) as the outgroup. In the resultant tree (Fig. 1), BKV sequences were divided into three major clusters with high BPs (100%), corresponding to subtypes I, subtypes II and III, subtype IV, respectively. The cluster comprising subtypes II and III was divided further into two subclusters with high BPs (100%); subtype I was subdivided into four subclusters with high BPs (100%), corresponding to a, b-1, b-2 and c (39); and subtype IV was divided into six subclusters with high BPs (93–100%), corresponding to a-1, a-2, b-1, b-2, c-1 and c-2 (Nishimoto et al., 2007). In addition, three pairs of subtype IV subclusters (a-1 and a-2, b-1 and b-2, and c-1 and c-2) showed sister relationships with higher BPs (66–92%). Table 1 shows the geographical and phylogenetic distribution of the 145 complete BKV genomes.

**Integrity of coding sequences contained in complete BKV genomes**

We aligned the entire coding sequences of the 145 complete BKV genomes (Supplementary Table S2). This alignment revealed that in most complete genomes all the genes encoding six viral proteins (VP1–3, agnoprotein, large T and small t) had neither a gap (i.e. a deletion or insertion) causing a frame shift in transcription nor a nucleotide substitution causing immature termination. However, two types of gaps occurred in certain BKV genomes. First, some BKV genomes lacked one of two 9 nt regions within the agnoprotein gene [nt 514–522 or 550–558; the nucleotide numbering is that of Dunlop (abbreviated DUN) (GenBank accession no. V01108)]. The nt 514–522 deletion was found in only a single genome (ITA-5), whereas the nt 550–558 deletion was present in many genomes, including two of three subtype II genomes, all five genomes of subtype IV subgroup a-2 and all eight genomes of subtype IV subgroup b-2 (Table 2). Second, all the subtype II and III genomes lacked a 6 nt region (nt 2884–2799) and all three subtype III genomes lacked another 6 nt region (nt 2820–2815); both these regions were located in the 3’-terminal region of the large T antigen gene (Table 2). None of the gaps resulted in a frame shift and most were unique to certain subtypes or subgroups of BKV (Table 2), suggesting that most gaps detected in the coding region of the genome were generated during branching of BKV into subtypes and subgroups and excluded the possibility that they occurred accidentally during the analysis.

**Archetypal TCRs unique to subtypes and subgroups**

Alignment of all TCR sequences of the 145 complete genomes (Supplementary Table S2) revealed that only four (two subtype I and two subtype IV genomes) had rearrangements involving long sequences (the structures of these rearranged TCRs will be described in the next section). The incidence of rearranged TCRs was 2/85 (2.4%) for subtype I and 2/58 (3.4%) for subtype IV, with no significant differences with the incidence of rearranged TCRs between subtypes I and IV. Consensus TCR sequences were identified for subtype I subgroups (a, b-1, b-2 and c), subtype II, subtype III, and related subtype IV subgroups (i.e. a-1 and a-2, b-1 and b-2, and c-1 and c-2) (Fig. 2). These consensus sequences (also designated subtype- and subgroup-specific TCRs) were used to construct a prototypic TCR sequence with nucleotides common to all or most subtypes and subgroups at all positions (Fig. 2). Each subtype- or subgroup-specific TCR had neither deletion nor duplication of long nucleotide stretches (2 bp or longer for a deletion and 5 bp or longer for a duplication) throughout the TCRs (Fig. 2). However, they were distinguished by the presence of some nucleotide mismatches and single-nucleotide deletions; the latter occurred at two sites (nt 92 and 254) in the subtype II and III TCRs and in the subtype I TCR, respectively (Fig. 2). Therefore, subtype- and subgroup-specific TCRs may have been generated from the prototypic TCR in the course of BKV evolution, while the archetypal configuration has been conserved.

**Structures of rearranged TCRs**

Rearranged TCRs were identified in two subtype I (MMR-6 and NAR-16) and two subtype IV (SEC-21 and NEA-28) genomes. The structures of these rearranged TCRs are depicted in Fig. 3, with four blocks (P, Q, R and S) forming the archetypal TCR (Moens & Van Gheluwe, 2005) shown by rectangles. The MMR-6 and NAR-16 TCRs had short duplications, while NEA-28 had a long duplication. The
Fig. 1. NJ phylogenetic tree classifying 145 complete BKV genomes into subtypes and subgroups. Complete genome sequences, excluding the TCR sequences, determined in this and previous studies (Supplementary Table S2), together with reference sequences [AS, DUN, JL, KOM-7, MT, RYU-3, TW-3 and WW (Seif et al., 1979; Tavis et al., 1989; Nishimoto et al., 2006)], were used to reconstruct the phylogenetic tree. The baboon polyomavirus SA12 (GenBank accession no. AY614708) was used as the outgroup. The phylogenetic tree was visualized using the NJPlot program. Subtypes and subgroups are indicated to the right of the tree, and the numbers at nodes are BPs (%) obtained for 1000 replicates (only those for major clusters are shown). Bar, 0.02 substitutions per site.
duplicated segment (21 bp) in the MMR-6 TCR spanned nt 101–121 and mainly included an origin-proximal region of the Q block, with duplication of only a 2 bp sequence at the origin-distal end of the P block; the duplicated segment (15 bp) in the NAR-16 TCR spanned nt 138–151 and mainly included an origin-proximal region of the R block; and the duplicated segment (103 bp) in NEA-28 spanned nt 169–271 and included an origin-distal region, the entire S block and the initiation site of the agnoprotein (Fig. 3).

The SEC-21 TCR had a deletion of a short segment (nt 150–119) located within the R block. These features of the rearranged TCRs are unique, in marked contrast to those detected previously in clinical samples from immunocompromised patients, since the latter usually involved complete or partial duplication of the P block (Moens & Van Ghelue, 2005).

Phylogenetic relationships of BKV genomes with and without TCR rearrangements

In the phylogenetic tree (Fig. 1), genomes with rearrangements (MMR-6, NAR-16, SEC-21 and NEA-28) are shown with white letters on black backgrounds. MMR-6 and NAR-16 are included in subclusters corresponding to the b-1 and c subgroups of subtype I, respectively, and SEC-21 and NEA-28 are included in subclusters corresponding to the a-1 and b-1 subgroups of subtype IV, respectively (Fig. 1). Thus, it can be concluded that the genomes with TCR rearrangements do not constitute an independent cluster distinguishable from those without TCR rearrangements. This result suggests that the BKV genomes with rearranged TCRs were generated during infection of individual hosts.

DISCUSSION

This study was undertaken to examine the archetype hypothesis from a phylogenetic viewpoint. Instead of amplifying subgenomic fragments, we amplified and sequenced complete viral genomes from urine samples of non-immunocompromised individuals. This approach allowed us not only to characterize TCR structures, but also to conduct an unambiguous phylogenetic analysis of BKV isolates based on complete genomes [it was unlikely that PCR preferentially amplified genomes with either archetypal or rearranged TCRs, since PCR has been used widely to detect both archetypal and rearranged BKV TCRs (Moens & Van Ghelue, 2005)]. Moreover, our approach was useful in examining whether the detected TCRs were

Table 1. Distribution of BKV genomes analysed in this study

<table>
<thead>
<tr>
<th>Geographical area</th>
<th>Subtype I</th>
<th>Subtype II</th>
<th>Subtype III</th>
<th>Subtype IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subgroup</td>
<td>Total</td>
<td>Subgroup</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>b-1</td>
<td>b-2</td>
<td>c</td>
</tr>
<tr>
<td>Europe</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Africa</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>W Asia</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NE Asia</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>SE Asia</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>USA</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>14</td>
<td>22</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 2. Deletions detected in the coding region of the BKV genome

<table>
<thead>
<tr>
<th>Gene</th>
<th>Deleted nucleotide stretches*</th>
<th>Changes in amino acid sequences</th>
<th>Isolates carrying deletions†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agogene</td>
<td>514–522</td>
<td>Deletion 43–45</td>
<td>MMR-29 (IV/c-1)</td>
</tr>
<tr>
<td>Large T antigen</td>
<td>2820–2815</td>
<td>Deletion 670–671; substitution of 669N for 669R</td>
<td>FUK-22, NEA-27 (III)</td>
</tr>
</tbody>
</table>

* Nucleotide numbers are those of Dunlop (GenBank accession no. V01108).
† Subtypes and subgroups are indicated within parentheses.
derived from intact viruses. The findings demonstrate that the archetypal TCR configuration has been conserved during the evolution of BKV.

Virological evidence for the archetypal hypothesis can be found in the literature. Sugimoto et al. (1989) reported a case of systemic lupus erythematosus in which various complete BKV genomes with archetypal as well as rearranged TCRs were directly cloned from the urine of the patient. Interestingly, about half of the recombinant clones had a very rare mutation (BglI, a mutation generating a novel BglI cleavage site at the origin of replication). In this case, there seems to be no doubt that a BglI clone carrying an archetypal TCR generated BglI clones carrying various rearranged TCRs. Furthermore, Nukuzuma et al. (2006) propagated archetypal BKV strains belonging to subtype I in human renal proximal tubule epithelial cells (HPTE cells), which are possible target cells in productive BKV infection in vivo (Randhawa et al., 1999). The archetypal configuration of the TCRs was essentially conserved during viral replication in HPTE cells, but variants carrying rearranged TCRs with duplications or deletions emerged occasionally (Nukuzuma et al., 2006). This observation mimics our findings that BKV isolates from non-immunocompromised individuals usually carried archetypal TCRs, with rare occurrence of rearranged TCRs. Therefore, the phylogenetic evidence together with the virological observations (Sugimoto et al., 1989; Nukuzuma et al., 2006) provide substantial support for the archetypal hypothesis.

The TCR of a strain (WW) belonging to subtype I/subgroup b-1 is commonly used to represent the archetypal TCR of BKV (Knowles, 2001). Nevertheless, the present findings suggest that, like other parts of the genome, the BKV TCR underwent evolutionary changes involving nucleotide substitutions and single-nucleotide deletions. Thus, the TCR of each subtype or subgroup of BKV has a unique set of nucleotide substitutions and deletions. We therefore suggest that ‘archetype’ should be used as a conceptual term that denotes a prototypical structure that can generate various rearranged TCRs during viral growth in vivo and in vitro. In this sense, the consensus TCRs for respective subtypes or subgroups (Fig. 2) are all archetypal.

To date, archetypal TCRs (or ‘WW-like’ TCRs) with various nucleotide substitutions and single-nucleotide insertions (or deletions) have been identified in several geographical regions (reviewed by Moens & Van Ghelu, 2005). Upon closer inspection of these changes, many of them agreed with those shown in Fig. 2. For example, the WWT TCR detected in Norwegian children was thought to be a variation of the WW TCR with a few nucleotide mismatches (Sundsfjord et al., 1990), but this TCR was identical to the I/b-2 TCR (Fig. 2). Since subtype I/subgroup b-2 is prevalent among Europeans (Ikegaya et al., 2006; Zheng et al., 2007), nucleotide substitutions unique to the I/b-2 TCR (Fig. 2) have frequently been identified in TCRs from Europeans (Moens & Van Ghelu, 2005). In short, archetypal TCRs detected in clinical samples should be evaluated with reference to all the subtype- and subgroup-specific TCRs shown in Fig. 2.

Fig. 2. Archetypal TCR sequences detected from the urine of non-immunocompromised individuals. Sequences between the midpoint of the origin of replication and the start site of the agnoprotein gene are shown. Consensus TCR sequences were identified for subtype I subgroups (a, b-1, b-2 and c), subtype II, subtype III and related subtype IV subgroups (i.e. a-1 and a-2, b-1 and b-2, and c-1 and c-2), and then a prototypic (Proto.) TCR sequence with nucleotides common to all or most of these consensus sequences was identified. The prototypic TCR is shown at the top of the figure and the consensus sequences for subtype I subgroups, subtype II, subtype III and subtype IV subgroups are shown below in relation to the prototypic TCR, with similar nucleotides indicated by dots and deletions identified by rectangles. IV/a, IV/b and IV/c indicate the consensus TCRs for subgroups a1 and a2, b1 and b2, c1 and c2, respectively, of subtype IV. Blocks (P, Q, R and S) commonly used to denote archetypal TCRs (Moens & Van Ghelu, 2005) are indicated above the prototypic TCR sequence.
Finally, our results provide basic information on the rearrangement of the BKV TCR in non-immunocompromised individuals. First, we found in a large dataset that the incidence of rearranged BKV TCRs in the urine of such individuals is low (2–4 %), in contrast to the surprisingly high percentage (22/47, 46.8 %) of rearranged TCRs found previously in tissues from healthy subjects (based on pooling of data from independent studies) (Sharma et al., 2007). Second, the detection rates for rearranged TCRs were 2.4 and 3.4 % for subtypes I and IV, respectively, with no significant difference between these detection rates. Third, the rearranged TCRs in the urine samples of immunocompetent subjects were unique in that they do not have significant duplication of the P block, which is frequent in TCRs (e.g. T2R, TC-3, BKVAN-1 and -2) from the urine and kidneys of renal transplant patients and in those (e.g. URO1, PBMC-9b, -18 and -21) from other cellular sources (Moens & Van Ghelue, 2005).

ACKNOWLEDGEMENTS

We are grateful to all the donors of the urine samples. This study was supported in part by grants from the Ministry of Health, Labour and Welfare, Japan.

REFERENCES


