The complete genome sequence BK polyomavirus study in kidney transplanted patients

Manoochehr Makvandi, Gholam Abbas Kaydani, Heshmatollah Shahbazian, Alireza Teimoori, Fakher Rahim, Sajjad Aslani

1Department of Virology, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
2Golestan Hospital, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
3Health Research Institute, Thalassemia and Hemoglobinopathies Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
4Department of Microbiology, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

ABSTRACT

Background: BK polyomavirus is a member of the Polyomaviridae. This virus has spread worldwide and up to 82% of the world populations are serologically positive. BK polyomavirus usually transmits through inhalation or fecal-oral way in childhood, as well as is likely to cause an asymptomatic disease. The virus can be reactivated in people who are immunocompromised.

Objectives: In this study we aimed to determine the complete genome sequence of BK polyomavirus in 5 patients receiving kidney transplantation in Golestan hospital, Ahvaz, south west of Iran.

Materials and Methods: The study was performed on urine and blood samples from kidney transplant ward, Golestan hospital, Ahvaz, Iran. To amplify the whole genome of BK polyomavirus AccuPower ProFi Taq PCR PreMix kit (Bioneer, South Korea) was used. Then, purified fragments were cloned into the vector. Sequences derived from the sequencing process were assembled by MEGA7 software. Data were analyzed using MEGA7 software and a phylogenetic tree was built based on the maximum likelihood ratio method.

Results: Overall, 40 urine samples (40%) and two plasma samples (12%) were positive for BK virus DNA. For the selected samples, W.G. Long PCR was performed and 5.1 kbp fragment was observed in all five samples. These fragments were cloned and sequenced. Genomic sequence analysis of the 5 strains studied showed 97.6% homology and all our study samples were of the same clade, which could be a reason for our patients to be infected with the same strain. Comparing the 5 isolates in our study with reference strains, showed more than 98% of homology, and variation was observed in less than 2% of nucleotides.

Conclusion: These five isolates showed more than 98% homology compared to the reference sequences, which can be attributed to the fact that these patients are infected with a common strain or that the genomic stability in the strain exists in our geographic area. We can conclude that in our geographic region there is a genomic stability and this strain can also be used as a positive control sample in other parts of the country.

Implication for health policy/practice/research/medical education:
The event of BK virus nephropathy practically in kidney transplant recipients shows that a merging of factors present around local injury and immunosuppression may lead to several complications, adverse events, and ultimately decreasing graft survival. Complete-genome sequencing leads to the availability of genomic information that can be used as a key to understanding the causes of health and disease. This is not a simple or rapid duty and will need considerable investment and personal, institutional, and governmental support; but with time and efforts, this technique can undoubtedly help in rapid examination of the etiology for complex diseases.


*Corresponding author: Gholam Abbas Kaydani, Email: kайдani56g@gmail.com
1. Background

BK polyomavirus is a member of the Polyomaviridae with a circular genome around 5.13 kbp in size. This is a non-enveloped double-stranded DNA virus with an icosahedral capsid composed of the three main proteins, including VP1, VP2, VP3, of which VP1 is the most important one (1,2).

This virus has spread worldwide and up to 82% of the world populations are serologically positive (3). BK polyomavirus usually transmits through inhalation or fecal-oral way in childhood, as well as is likely to cause an asymptomatic disease (4,5). Moreover, this virus remains hidden in the urinary tract, kidneys and sometimes mononuclear blood cells in the number of infected people. The virus sometimes is disposed of urine of healthy individuals (6). The virus can be found in people who are immunocompromised, such as people who have received organ transplants (especially kidney transplants), as well as pregnant women or people with AIDS, and will again be activated and causes illness. The virus has been isolated from urine of 10%-60% of people who have received kidney transplants and can develop nephropathy in some of these people (3,5,7-11). The virus can cause hemorrhagic cystitis or urinary tract obstruction in people who have a bone marrow transplant (7,12-15).

BK polyomavirus genome comprised of double-stranded DNA, and be partitioned into three parts, including early, late and non-coding regions. The early region encodes the large T antigen (TAg), the small t antigens (tAg) and the short t antigen proteins. Hence the late region encodes VP1, VP2, VP3 structural proteins and a non-structural protein called agnoprotein (16-23). BK polyomavirus can also be classified based on non-genetic and serological testing into several genotypes. Given the impact of the restriction endonuclease on the gene fragment VP1, it can be sub-classified into four types, including virus sub-type 1, 2, 3, and 4, which this classification is fully compatible with their serology. By mutations in the gene sequence of VP1, subtypes 1 and 4 can be further divided into sub-groups, of which the sub-type 1 is classified into four subgroups (1/a, 1/b-1, 1/b-2, 1/c) and subtype 4 is classified into six subgroups (4/a-1, 4/a-2, 4/b-1, 4/b-2, 4/c-1, 4/c-2). The subtypes 1 have spread around the world, while subtype 4 can be seen only in East Asia. Moreover, subtypes 2 and 3 have been rarely reported in some areas (9,11,24-29).

Usually multiple mutations occur in the VP1 gene, but sometimes gene mutations also occur in the TAg. The Tag protein is the most important regulatory protein of the virus, which plays an important role in the virus life cycle, and is required for BK virus genome replication and activation of the switching the genes amplification from early to late genes. Loci polymorphisms in the noncoding control region (NCCR) genome of the virus has been reported. Substitutions and changes in some specific nucleotides in the genome of the virus can alter the amino acid residues in the protein core, and this change may play an important role in the pathogenesis, characteristics, and survival of the virus (16-23).

Few studies that examined the entire genome of BK polyomavirus have been done in the world. The most important strains of BK polyomavirus, including MM, DUN and AS have been reported by Chen et al (30). To the best of our knowledge, no study has not performed regarding the complete genome BK polyomavirus in Iran.

2. Objectives

The aim of this study was to determine the complete genome sequence of BK polyomavirus in five patients receiving kidney transplantation in Golestan hospital, Ahvaz, Iran.

3. Materials and Methods

3.1. Sample collection and DNA extraction

The study was performed on urine and blood samples from kidney transplant ward, Golestan hospital, Ahvaz, Iran. A total of 100 urine and 20 plasma samples were selected. Polymerase chain reaction (PCR) test for the presence of BK polyomavirus genome was performed for all of the samples. To collect urinary epithelial cells, samples were centrifuged at 3000 rpm and sediments were kept at -70°C. BK polyomavirus genome extraction from urine and plasma samples was done using the high pure viral nucleic acid kit (Roche, Germany) based on the manufacturer’s instructions.

3.2. PCR testing to determine the BKV positive samples

Conventional PCR method with specific VP1 primers was used to determine the existence of BKV genome. Around 100 ng of extracted DNA from clinical specimens was used in 50 μL reaction including (Amplicon) Master Mix and 0.5 μM of each of specific primers for VP1 to find 300 bp band. Because of the possibility of contamination with JC and SV40 polyomavirus, conventional PCR with primers specific for these viruses was performed on extracted DNA and in case of the existence of the band 610 bp for JC and bands 143 bp to SV40, the sample was excluded. A temperature program for amplicons of these components performed by the thermocycler is as follows (Table 1).

Then, all amplified products were analyzed by electrophoresis on 1.5% agarose gel. Of those samples, four samples of urine and one plasma sample were selected and enrolled in the study.
### 3.3. Long PCR test for BKV genome

To amplify the whole genome of BK polyomavirus, AccuPower ProFi Taq PCR PreMix kit (Bioneer, South Korea) was used. DNA extracted from urine and plasma samples was digested the Bam HI enzyme. In this reaction, 6 μl buffer, 1 unit of Bam HI enzyme was added to 3 μl of DNA and incubated for 1 hour at 37°C. Then, 5 μl of the digested DNA was transferred in 50 μl PCR reaction with 0.5 mM of each specific primers. Then, 5.1 kbp fragments were analyzed by electrophoresis on 1% agarose gel. A temperature program used for the long PCR (Table 2).

### 3.4. Complete genome of BKV cloning

After seeing the 5.1 kbp fragments by gel electrophoresis, the final product was purified by GF-1 PCR Clean-up kit (Vivantis Inc., Malaysia). Then, purified fragments were cloned into the vector by InsTAclone PCR Cloning kit (Thermo Fisher Scientific, Lithuania) according to manufacturer’s instructions. Next, the products were transformed into the DH5α bacteria. After completion of the culture and validates transformation, a single colony was selected and its plasmid was extracted by a high pure plasmid isolation kit (Roche Inc., Mannheim, Germany) according to manufacturer's instructions. To verify the integrity of the fragments of interest, PCR was performed by specific Bam HI primers for each fragment. The observation of 493 bp fragment indicates the accuracy of the work. The sequences included in TA vector was sequenced using seven primers (Table 3) and both direction sequences were obtained. The sequencing was done using BigDye solution in an ABI Biosystems™ sequencer.

### 3.5. BK virus genome analysis software

Sequences derived from the sequencing process was assembled by MEGA6 software based on the strain of BK polyomavirus isolate BKV34-2 or Dunlop (KP412983.1), and overlapping areas were excluded. Each of the completed sequences was compared with reference sequences using MEGA7 software. Moreover, sequences of VP1, VP2, VP3, LTag, stAg and Agno regions of each sample were compared with homologous sequences in the Gene bank (AB263936.1/AB369087.1/FR720311.1/FR720318.1/JN793994.1/ JN794026.1/V01108.1/V01109) and analyzed using MEGA7 software and phylogenetic tree was built based on the Maximum likelihood ratio method. Tree of bootstrapping method was used for reliability. Nucleotide substitution model was done according to Hasegawa-Kishino-Yano. Genomic Map of isolates of interest has drawn by SnapGene software.

### 4. Results

#### 4.1. Qualitative PCR

One hundred urine specimens and 20 plasma samples of kidney transplant recipients were collected from the laboratory, and PCR was performed for the genome of BK virus. Overall, 40 urine samples (40%) and two plasma samples (12%) were positive for BK virus DNA. These specimens were positive for the presence of a 300 bp VP1 gene for the BK virus. For all positive samples, PCR test was performed with specific viral JCV and SV40 primers. All samples were negative for SV40 genome and 19 samples (45.2%) were positive for JCV, all of which were urine specimens.

#### 4.2. Whole genome long PCR

For the selected samples, W.G. Long PCR was performed and 5.1 kbp fragment was observed in all five

---

### Table 1. JCV and SV40 amplicons temperature program

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation denaturation</td>
<td>95°C</td>
<td>2 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>5 min</td>
</tr>
</tbody>
</table>

### Table 2. A temperature program used for the long PCR

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation denaturation</td>
<td>98°C</td>
<td>5 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>10 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C</td>
<td>45 s</td>
</tr>
<tr>
<td>Extension</td>
<td>68°C</td>
<td>7.5 min</td>
</tr>
<tr>
<td>Final extension</td>
<td>68°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>

### Table 3. Primers used in the Long PCR and sequencing

<table>
<thead>
<tr>
<th>Size</th>
<th>Sequence (5’→3’)</th>
<th>Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>879bp</td>
<td>CTAGGAATCTGGGCTTTGTCG</td>
<td>BK345F</td>
</tr>
<tr>
<td>998bp</td>
<td>GTAAATTCGTGTCGCTGGGCCC</td>
<td>BK831F</td>
</tr>
<tr>
<td>784bp</td>
<td>GGGGATCCATCTGGGCTTTAG</td>
<td>BK1731F</td>
</tr>
<tr>
<td>1426bp</td>
<td>TGCGAGCATCCTGCCCAGTTTTT</td>
<td>BK2415R</td>
</tr>
<tr>
<td>866bp</td>
<td>CATCATTGTAACAGGCTGTCG</td>
<td>BK3216F</td>
</tr>
<tr>
<td>746bp</td>
<td>TCTTCTGTTAGCGTCTCTGGG</td>
<td>BK3712F</td>
</tr>
<tr>
<td>1700bp</td>
<td>GTATGTCATGATCATCTAGGGG</td>
<td>BK4076R</td>
</tr>
<tr>
<td>5.13 kbp</td>
<td>GGATCCCCAGATCCAGACTGCTC</td>
<td>BK1739R</td>
</tr>
</tbody>
</table>

(Continued on the next page)
samples. These fragments were cloned and sequenced. In analyzing the obtained sequences and comparing them with the reference sequences in the GenBank, the changes were observed as follows. In the agarose gel, the band observed from the plasma sample and two of the urine specimens were poorly presented.

4.3. Phylogenetic analysis and genotype determination of BKV sequences
The sequences from the five patients selected in this study were compared with the three reference sequences found in the GenBank that were obtained from previous studies. Genomic sequence analysis of the 5 strains studied showed 97.6% homology and all our study samples were of the same clade, which could be a reason for our patients to be infected with the same strain. Comparing the 5 isolates in our study with reference strains showed more than 98% of homology, and variation was observed in less than 2% of nucleotides. In the LTAg gene of our strains, more than 97% of the amino acid sequence was found to have a similarity with the strain found in GenBank. This similarity in the VP1, VP2, and VP3 genes was 99, 100, and 100%, respectively. Due to the relatively high homology of the isolated isolates, isolate AJUMS/Kaydani.25 BK virus was selected for further analysis. The genomic map of this isolate is shown in Figure 1. Based on the complete sequence obtained from BK virus, the following results were obtained.

4.4. Sample genotype and sub-types determination
As outlined above, the BK virus has four sub-types. This definition of various subtypes is based on changes in the nucleotide sequence of the VP1 gene. The complete sequence of VP1 gene was determined and compared to the sequences in the gene bank using MEGA7 software, and it was found that all five were subtypes I.

4.5. Analysis the changes in VP1 gene and its protein
Comparison of complete sequence of VP1 gene clinical specimen No. 25 with standard strains of Dunlop, MM and SJH-LG-253 showed that the nucleotide sequence 20, 20 and 9 nucleotides were different, respectively. Differences in amino acids were also found in amino acid 1, 6 and 1, respectively.

4.6. Study of changes in large T Ag, VP2, VP3, Agno and Small t antigens and analysis of their protein
As previously stated Large T Antigen is the most important regulatory protein in the virus and plays an important role in replicating the virus. In this study, the complete sequences of Large T Ag, VP2, VP3, Agno and small t antigen genes were aligned and compared to the standard strains of Dunlop, MM and SJH-LG-253. Nucleotide and amino acid differences are presented in Tables 4 and 5. Nucleotide changes have occurred on sites A or T, which cannot be random. According to the results obtained in the above tables, specimen No. 25 is genetically very close to the Dunlop strain followed by the SJH-LG-253 strain, and is different from the MM in some genes (Figure 2).

5. Discussion
Today, BKV infection is one of the most common infections in renal transplant recipients, and BK polyomavirus-associated nephropathy (BKVPN) is seen about 10% of patients who have had kidney transplants, which causes renal dysfunction and even rejection of the transplanted organ. BKV is common in serology and it is unclear whether a specific subtype or variant will play a role in its pathogenesis (1,6). Treatment for this infection has been difficult and there

<table>
<thead>
<tr>
<th>Table 4. Nucleotide difference between standard isolates and isolated strains No. 25</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJUMS.Kaydani-25 BK virus</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Dunlop</td>
</tr>
<tr>
<td>MM</td>
</tr>
<tr>
<td>SJH-LG-253</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5. Amino acids difference between standard isolates and isolated strains No. 25</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJUMS.Kaydani-25 BK virus</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Dunlop</td>
</tr>
<tr>
<td>MM</td>
</tr>
<tr>
<td>SJH-LG-253</td>
</tr>
</tbody>
</table>
has not yet been a proven medication to treat the disease associated with the virus, but recent research has shown that leflunomide, along with cidofovir, can greatly inhibit the proliferation and expression of BK virus genome (31).

In this study, we determined the sequence of the complete BK virus genome isolated from five patients who received kidney transplantation and were admitted to hospital due to fever and increased creatinine. These five isolates showed more than 98% homology compared to the reference sequences, which can be attributed to the fact that these patients are infected with a common strain or that the genomic stability in the strain exists in our geographic area.

Takasaka et al (15) also reported no changes in the nucleotide arrangement in two patients, and in the other three cases, the changes were negligible, resulting in a consistent record of the genome of the virus. Similarly, Chen et al (30) reported no changes in the isolated amino acids in their study, although there was a slight change in their nucleotide sequence. Burger-Calderon et al (8) and Henriksen et al (32,33) also showed very little point mutations in the isolates of interest.

Although the sequences of different genes in the whole genome of the virus can be sequenced individually and compared to other sequences in the GenBank, the determination of the entire sequence gives us more accurate results. Luo et al (25) found that partial amplification of the genome to determine subtypes and subgroups could not be decisive, and it would be better if the complete genome was sequenced. This result was confirmed in the present study. Comparing the different genes of VP1, VP2, VP3, LTAg, stAg, and Agno with reference strains of Dunlop, MM and SJH-LG-253, we found that the nucleotide difference was significant only in the LTAg gene and Agno, and had very little difference in the remaining genes. Therefore, in determining the polymorphism, it is better to use the VP1 sequence that is more stable. Seemayer et al (22) confirmed this finding and stated that the genetic study for determining the polymorphism on the VP1 sequence is much more suitable than LTAg and Agno. Boldorini et al (34) found that point mutations in the patient group were much higher than that of the healthy group, suggesting that nephropathy-related strains (BKVN) have less genetic stability. Sharma et al (28) also considered the major site for determining polymorphism as the VP1 gene, although they were significantly different in the LTAg region.

6. Conclusions

Therefore, we can conclude that in our geographic region there is a genomic stability and this strain can also be used as a positive control sample in other parts of the country.

Limitations of the study

This study was carried out for the first time in Iran. This study should be validated by larger sample size investigations.

Authors’ Contributions

GAK; study concept and design, and drafting of the manuscript. HS: clinical consultant and acquisition of data. SA; acquisition of data. MM; supervisor and analysis and interpretation of data. FR; critical revision of the manuscript for important intellectual content and statistical analysis. ART and GAK; administrative, technical, and material support. All authors reviewed and approved the final manuscript.

Conflicts of Interest

Authors declare no conflict of interests.

Funding/Support

This study was supported by a grant from the Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran (Grant No. B-9472)

References


Copyright © 2018 The Author(s); Published by Society of Diabetic Nephropathy Prevention. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.