UPPSALA
UNIVERSITET

Characterization of co-infections and minor variants of BK polyomavirus in clinical sample by NGS

Safia Khatoon

Master Degree Project in Infection Biology, 30 credits. Spring 2020
Department: Medical Biochemistry and Microbiology
Supervisor: Dr. Christina Öhrmalm, Dr. Anders Bergqvist

Abstract

BK polyomavirus (BKV) is associated with urinary apparatus pathogenesis in kidney transplant recipient. Immune suppression triggers BKV reactivation that potentially causes polyomavirus associated nephropathy (PVAN), a major post-transplant problem causes graft rejection. Antiviral immunity plays the key role in limiting the viral replication but selection by the immune system or antivirals may cause the evolvement of escape variants with higher fitness. Mutation in VP1, the major capsid protein can allow BKV to escape neutralizing antibodies. In an attempt to detect co-infection and minor variants, BKV VP1 genomic region was amplified by PCR and analysed by deep sequencing from plasma samples of four kidney transplant recipients. BKV genotype I and IV was identified in patients and each patient was detected with one BKV genotype. Multiple point mutations and subsequent changes in amino acid were detected in majority, three out of four, of the patients.

Keywords
BK polyomavirus, Co-infection, Mutation, Kidney transplant, Next generation sequencing

Popular summary

To many of us viruses are the stuff of mystery and danger. They can cause detrimental infections by knocking down our immune system or they may get knocked down by the immune system and not being able to do any harm other than causing some flu like symptoms, which is fact in most of the case. But there are viruses that enters our body, find a way to hide from immune system and stay forever within the body, not doing any harm as because immune system is in control. What if the immune system is not at its best due to a condition like organ transplantation? Readily the hidden virus takes over the battle! BK polyomavirus (BKV) is one of those viruses, which enters the body early in life and stays lifelong in kidney epithelial cells. When a person is under immune suppression after kidney transplantation, to prevent transplanted organ loss, BKV takes control and start growing in numbers, causing severe diseases like severe urinary tract infection, graft dysfunction, and polyomavirus associated nephropathy (PVAN). There is lack of effective antivirals against BKV and immune system can control viral reproduction, so withdrawal of immune suppression is so far best treatment option.

There are many types of BKV and there is possibility that kidney from a donor was infected with another type of BKV than the kidney transplant recipient, in that case recipient's immune system is not enough in controlling diseases condition. Moreover, recipient can also get co-infected with another BKV types from donor kidney. Virus can also find a way to escape from immune system through gaining changes in their genes. For this study we set out to investigate if a patient is infected with more than one types of BKV or if the virus that infected patient gained any changes to escape the immune system. We have identified that each of the patients studied was infected with one type of BKV. We also have detected changes in viral gene, which may open new avenue for further investigation.

List of Abbreviations

BKV	BK Polyomavirus
PVAN	Polyomavirus Associated Nephropathy
KTRs	Kidney Transplant Recipients
miRNA	MicroRNA
NGS	Next Generation Sequencing
Pt	Patient

List of content

Introduction …….. 6
Aim ……... 10
Material and Methods ……... 11
Patient Sample …….. 11
Ethical Statement …….. 11
DNA Extraction ……..... 11
Sample Preparation Strategies …….. 11
BKV VP1 PCR ……. 12
Nextera XT DNA Library Preparation ……..… 13
MiSeq Dual Indexing Sample Preparation ……... 13
Gel Electrophoresis …….. 14
Next Generation Sequencing Analysis ……. 14
Results …….. 16
BK VP1 Nested PCR …… 16
Next generation sequencing ……. 17
Phylogenetic analysis of BKV VP1 Sequences ……... 18
Mutation analysis of BKV VP1 Sequences ……. 19
MiSeq amplicon sequencing …… 20
Discussion …… 22
Acknowledgement …… 24
References …… 25

Introduction

BK polyomavirus (BKV) and adenovirus are well recognized disease-causing agents, especially for viral infections in renal transplant recipient, worldwide because of their vast range of distribution. These DNA viruses can establish persistent infections and are able to reactivate after infected individual receive immunosuppressive therapy due to solid organ transplantation. In kidney transplant recipient (KTRs), the viruses are linked to systemic diseases or graft disfunction (1).

BKV was first isolated in 1971 by Gardner *et al*. from the urine of an immunosuppressed renal allograft recipient with ureteric stenosis in (2). BK polyomavirus is widely omnipresent in all human populations. In most people, BK polyomavirus established latent infection in the early childhood, and approximately 70 % of the children get infected by the age of ten (3,4). Usually in immunocompetent individual the primary infection does not cause any symptomatic disease, however, it can establish latent infection in renal tubular cell or uroepithelial cell for the rest of the life (5). Notably, at least 75 % adults are seropositive for BKV. Nevertheless, in immunocompromised individual or in patient whose immune system is suppressed, either due to organ (kidney) transplantation or by hematopoietic stem cell transplantation, the latent infection can be reactivated and lead to severe diseases conditions such as severe urinary tract infection, graft dysfunction, and polyomavirus associated nephropathy (PVAN), graft loss in renal transplant recipients (6,7).

BKV, a member of beta polyomavirus genus of the *Polyomaviridae* family, is a double stranded, nonenveloped DNA virus with an icosahedral capsid of 40 to 44 by diameter. The capsid is composed of 72 capsomers, each consists of five molecules of major capsid protein. The minor capsid proteins are found to be localized at the inner face of the capsid (8).

The viral genome is a double-stranded circular DNA, which is packed with histone proteins and comprises three functional domains. The early region of the genome encodes for large T antigens and small T antigens whereas the late region encodes structural protein VP1, VP2, and VP3 (9). Among three of the viral capsid protein only VP1 is exposed on the exterior side of the virion, containing five external loop BC, DE, GH, EF and HI, and therefore is responsible for the viral attachment to host cell receptor to mediate viral entry into the host cell (10,11). The genome also comprises a noncoding control region needed for viral gene expression and virus replication (12).

Polyomavirus usually enter the host cell via clathrin-mediated entry but BKV enter the targeted cell using caveolae-mediated endocytosis pathway. The VP1 capsid protein bind to disialic acid motifs on gangliosides GD1b and GT1b receptors then enter the cell by caveolae-mediated endocytosis pathway to get into endoplasmic reticulum using microtubules (13). Viral uncoating then takes place via endocytic pathway followed by retranslocation of the virus to cytosol in order to achieve viral entry into the nucleus (14). After the second rearrangement of capsid protein, VP2 and VP3 facilitate viral entry into the nucleus (15). Viral replication highly depended on the transcriptional and DNA replicational machinery provided by growing host cell (6).

BKV strains are grouped into four genotypes (I, II, III and IV) based on short subtyping region present in the BC loop of VP1 major capsid protein (11). BKV genotype I is categorized further into four subtypes Ia, Ib1, Ib2 and Ic and BKV genotype IV is classified further into subtype IVa1, IVa2, IVb1, IVb2, IVc1 and IVc2. The genotype I is most and genotype IV is second most prevalent in human species whereas very few adults are found to be infected by genotype II and III (16,17). Extensive studies of worldwide distribution pattern show that genotype I is present in all over the world while genotype IV is geographically distributed towards Europe and east Asia. The other two genotypes, II and III, are seldom found worldwide (18).

The transmission of BKV typically occurs via respiratory route supported by the evidence of BKV infection in tonsillar tissue of young children (4). Seroprevalence studies show that the primary infection begins to occur after the disappearances of maternal antibody within the first few months after birth, however, about 10-30% of infants are already seropositive at this point (7). Other possible routes of transmission have also been suggested, such as fecal-oral, blood transfusion, and transplacental (19). The primary infection is speculated to be asymptomatic or a mild respiratory illness (20). However, after the primary infection the virus disseminate to the urinary tract, where it remains latent for lifetime (21). Occasional reactivation of latency results in viral shedding in urine of both healthy and immune compromised individual although periodic viruria is more common for non-immunocompetent individual in comparison to immunocompetent individual (22).

The mechanisms of latency and reactivation of BKV are still not well understood. Profound studies on the mechanism of latency for herpesviruses suggested that viral lytic replication is controlled by the regulation of various genes (19). MicroRNAs (miRNAs) encoded by herpesviruses are found to act as regulators of viral replication. Similarly, to those of herpesvirus BKV also encodes miRNAs and evidence showed that BKV replication are controlled by high level of expression of miRNA (23). miRNA-dependent autoregulation of viral replication have also been demonstrated (24).

The reactivation of BKV from latency can be serious or fatal in immunocompromised individual. The foremost clinical manifestations of BKV infection are PVAN, ureteric stricture, tubulointerstitial nephritis and premature graft failure, while other less frequent clinical manifestations are pneumonitis, hepatitis, retinitis, and meningoencephalitis. In bone marrow transplant recipient, BKV reported to cause hemorrahagic cystitis (25–27). Under the circumstance of immunosuppression, BKV can reactivate from latency in the renal tubular epithelial cell, where viruses replicate up to a pathogenic level, causes necrosis, lytic distraction and denudation of the basement membrane leading to PVAN in kidney transplant recipient. Viruria and viremia resulting from viral replication are found in 80% of renal transplant patient and 10% of the progress to PVAN (22,28,29). BKV has been reported to reactivate soon after transplantation and seen in patient by about three months post-transplantation (22).

There is no single risk factor for the development of PVAN. Beside the immunosuppression, the other risk factors for BKV infection may include prior tubular injury from rejection, seronegativity of the recipient, older age, greater HLA mismatch, and diabetes mellitus (25).

Evidence suggests that the infection in the donor kidney elevates the risk of BKV associated diseases development after transplantation, specifically when donor and recipient harbor different BKV genotype (30). Additionally, one study shows that virus in the donor kidney can replicate to pathogenic levels in a recipient who primarily does not have neutralizing antibody responses to the viral genotype present in the donor kidney (31).

In solid organ transplant or in bone marrow transplant, a lack of immunocompetence due to immune suppressive therapy not only increase the risk of viral reactivation in transplant patient but also raise the risk of co-infection with two or more pathogen (1). Other than risk of co-infection with multiple virus, co-infection with different genotype of same viral species are also possible. Indeed, two different BKV genotype I (subtype Ib2) and IV co-infecting a single hematopoietic stem cell transplant recipient has been reported by Van *et al*. (32). One study conducted by Peretti *e.t al.* (31) showed that donor kidney which is co-infected with BKV genotype I and genotype IV can efficiently neutralize BKV genotype I and IV and get transferred to recipient via graft. Recipients who are initially infected with BKV type I can acquired BKV type IV infection from donor.

Despite the fact that there is no homogeneity of risk factors, active viral replication appears as the only common features to all KTRs at the risk of PAVN. Thus, timely screening of BKV replication may allow earlier intervention. Although the diagnosis of PVAN earlier require a biopsy it can also be performed by the detection of viral replication in the urine, plasma or cerebrospinal fluid by PCR based quantification. Likely, PCR have high sensitivity for detecting the active viral replication, which can also be detected by urine mRNA PCR (33).

In spite of the growing clinical requirement, virus specific antiviral agents for BKV are still unavailable and since antiviral immunity play the key role in limiting viral replication, reduction of immunosuppression is the only effective treatment as per evidence (19). Regular screening of KTRs to detect presence of BKV in blood stream and timely reduction of immunosuppression improves immunity (34). Although BKV infection acquired in early childhood induces neutralizing antibody response specifically against BKV and play a major role in neutralizing circulating virus, existence of BKV-specific antibody alone is not sufficient to provide protection against viral reactivation and associated diseases (35). Moreover, mutation in the viral capsid antigen allow virus escape from antibody mediated neutralization. One study showed that due to the variation in antibody receptor binding, some BKV genotype are able to escape antibody mediated neutralization raised against another genotype (36). Evidence supported that APOBC3 protein, which provide innate immunity against viruses by drastically mutating the genome of DNA viruses, is used by BKV to achieve beneficial mutations, which in turn allow the virus to escape antibody mediated neutralization (31). Significantly frequent base pair mutation were observed in the VP1 region of BKV isolated from patient with PVAN and it is speculated that VP1 nucleotide and amino acid variation possibly plays a role behind BKV escaping host immunity and antiviral resistance (37).

Infections by DNA and RNA viruses that are associated with viral persistence often display a quite complex pattern. First, due to genotypic variation they can cause co-infections. Second, due to evolutionary pressure by antivirals or the immune system, escape variants with higher

fitness might evolve. In this study we performed deep sequencing to detect coinfection and minor variants in VP1 genomic region of BKV present in the plasma sample of kidney transplant recipient that are not readily detected by conventional typing procedure such as Sanger sequencing. Sanger sequencing has been standard for sequencing of individual target genes but provides less data outcome (38). We used Next Generation Sequencing (NGS), a novel technique for deep sequencing analysis of the viral genome, to obtain more extensive characterization of the infection panorama. In comparison to Sanger sequencing, the NGS provides more data outcome. Additionally, NGS delivers highly quantitative analysis, since the starting sample is processed without cloning of single colonies. Moreover, it has faster turnaround and provide tremendously low sequencing cost per nucleotide (39).

Aim of the study

The aim of this project was to develop a procedure for the detection and identification of viral co-infections and minor variants in clinical samples. As a model system for the evaluation BK polyomavirus was used.

Materials and Methods

Patient Sample

For the deep sequencing analysis of the viral genome, plasma samples from four non-related BKV positive patients were used. Samples were collected from patient-1 at the day 124th, patient-2 at the day 745th, patient-3 at the day 107th, and patient-4 at the day 99th after kidney transplantation. The viral load in sample from patient 1- 4 were measured to be 82 000 copies/ml, 230 000 000 copies/ml, 82 000 copies/ml, and 99 000 copies/ml, respectively; prior to this study, and were deposited in Uppsala Biobank.

Ethical Statement

Clinical samples that were used have been consensually collected and deposited in Uppsala Biobank were used in accordance with local ethical guidelines and the Regional Ethical Review Board in Uppsala, Sweden (2015/488).

DNA Extraction

DNA was extracted from 200 µl of plasma using automated NucliSens EasyMag total nucleic acid extractor (bioMérieux, France), using the standard protocol for extracting both RNA and DNA. A total of 60 µl DNA was extracted.

Sample Preparation Strategies

Purified DNA from patients' sample was amplified using VP1 nested PCR. The first round PCR (outer PCR) was performed to amplify the VP1 gene of interest, and the inner portion of the gene was further amplified using a second round PCR (inner PCR).

After the first round of PCR, two different strategies were employed to amplify first PCR amplicons (**Figure 1**). In the first strategy, amplicon from the first PCR was amplified using inner primer set (**Table 1**), which was then used to generate a tagmented library by using Nextera XT DNA Library Preparation Protocol. As in the in the first strategy, amplicon from first PCR was also used in the second strategy and subjected to a tagged PCR using virus specific primers containing a unique set off index that were designed and analyzed using Primer Inspector software.

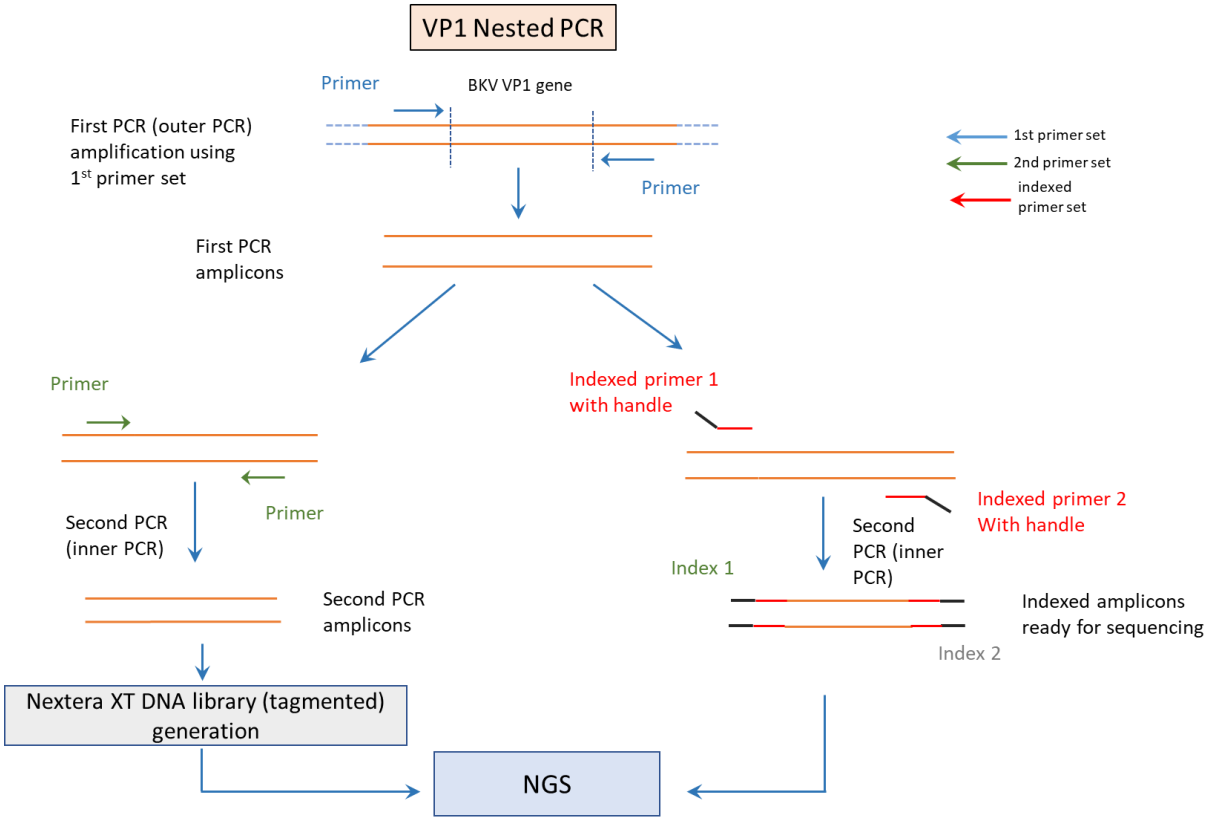

Figure 1: Illustration of methodological approach showing two different strategies of sample preparation leading to NGS.

BKV VP1 PCR

To amplify the genomic region corresponding to VP1 of BKV, a nested-PCR was performed with two different set of primers and conditions shown in **Table 1**. The amplifications were conducted in total volume of 25 µl containing 1 µl (10 µM) of each primer, forward and reverse, along with 12.5 µl of ready to use 2X Taq PCR master mix using manufacturer's instructions. In the first step, 5 µl template was added to the PCR mixture and in second step, 2µl of the amplified template was used. The reaction was carried out in Veriti 96-Well Thermal Cycler (Applied Biosystem). Agarose gel electrophoresis was carried out to check the fragment size of interest. The PCR product were cleaned up using ExoSapIT clean up kit (Thermo Scientific, Stockholm, Sweden) as per manufacturer's instructions to remove primers.

Table 1. Primers and amplification condition of BKV VP1 PCR.

PCR I (outer primers)	
Forward primer:	
BKFW1M	(5´-ACTATTG CCCCAGGAGG TGCTAATC-3´)
Reverse primer:	
BK-VP1-2RM	(5´-GBACTC CCTGCA TTTCC AAGGGG-3´)
PCR II (inner primers)	
Forward primer:	
BKFW2M	(5´-CCCGTG CAAGTGCCAAAACTRCT-3´)
Reverse primer:	
BKREV2M	(5´- CCTGCATGA AGGTTAA GCATRCTAG-3´)
Amplification condition: 94 ºC 4 min 35× (94 ºC 30 sec, 55 ºC 30 sec, 72 ºC 1 min) 72 ºC 5 min, Hold 4 ºC	

Nextera XT DNA Library Preparation

Before proceeding to Nextera XT DNA library generation (**Figure 1**, left panel), the cleaned amplified DNA from both rounds of nested PCR (eight products in total from 4 patients) were analyzed on 2200 Tape station (Agilent Technologies, Sweden AB) using D1000 Screen Tape Kit (5067-5583) following the manufacturer's protocol. However, for accurate quantification of input DNA, each DNA sample was subjected to Qubit dsDNA assay using Invitrogen Qubit 1X ds DNA HS Assay kit and Invitrogen Qubit 4 fluorometer (Thermo fisher Scientific, Stockholm, Sweden). For the library generation, the amplicons were tagmenteted, amplified, purified, pooled, and diluted to 0.2 ng/µl prior to sequencing using Nextera XT DNA library preparation Kit and following manufacturer's protocol. PCR amplification of libraries was performed using Nxtera PCR Mastermix, Index P7 primers and P5 adapter primers. The BKV VP1 region were sequenced using Illumina MiSeq sequencing (Illumina, San Diego, GA.).

MiSeq Duel Indexing Sample Preparation

The amplified product obtained from first round of previously mentioned VP1 nested PCR was in parallel utilized to make another indexed library (**Figure 1**, right panel). During the second nested round of PCR amplification a set of primer were designed containing unique index for each amplicon (**Table 2**) to generate desired barcoded amplicons. The amplifications were conducted in total volume of 25 µl containing 1 µl (10 µM) of each primer, forward and reverse, along with 12.5 µl of ready to use 2X Taq PCR master mix using manufacturer's instructions. Taq PCR master mix included Taq DNA polymerase

(QIAGEN), PCR buffer (QIAGEN), $MgCl_2$, dNTPs. Two µl of amplified template was added to 23 µl of reaction mixture and the reaction was carried out in a Veretti 96-Wall Thermal Cycler (Applied Biosystem). Gel electrophoresis was done to check the fragment size and PCR amplicons were cleaned up using ExoSapIT clean up kit (Thermo Scientific, Stockholm, Sweden) according to manufacturer's instructions.

Table 2. Primers and amplification conditions of tagged PCR with dual indexes.

PCR II inner primers	
Forward primer with index	
iF11v3BK	(5`-AATGATACGGCGACCACCGAGATCTACAC**gagcatgc** CCGTGCAAGTGCCAAAAC-3´)
iF12v3BK	(5´-AATGATACGGCGACCACCGAGATCTACAC**caatagcc** CCGTGCAAGTGCCAAAAC-3´)
iF13v3BK	(5´-AATGATACGGCGACCACCGAGATCTACAC**tgtcgatc** CCGTGCAAGTGCCAAAAC-3´)
iF15v3BK	(5´-AATGATACGGCGACCACCGAGATCTACAC**ctaatgcg** CCGTGCAAGTGCCAAAAC-3´)
Reverse primer with index	
iR11v3BK	3´-CAAGCAGAAGACGGCATACGAGAT**tcaggtgt** GTGACTGGAGTTCGCACTCCCTGCATTTCCAAG-5´
iR12v3BK	3´-CAAGCAGAAGACGGCATACGAGAT**gataagcc**GTGACTGGAGTTC GCACTCCCTGCATTTCCAAG-5´
iR13v3BK	3´-CAAGCAGAAGACGGCATACGAGAT**aaatcccg**GTGACTGGAGTTC GCACTCCCTGCATTTCCAAG-5´
iR15v3BK	3´-CAAGCAGAAGACGGCATACGAGAT**ttcgccgc**GTGACTGGAGTTC GCACTCCCTGCATTTCCAAG-5´
Amplification condition: 94 ºC 4 min 35 × (94 ºC 30 sec, 55 ºC 30 sec, 72 ºC 1 min) 72 ºC 5 min, Hold 4 ºC	

Gel Electrophoresis

To check amplified DNA fragment size, 8 µl of amplified product was mixed with 2 µl of 6× DNA loading dye and loaded on 1.8 % agarose gel. As size reference, 3 µl of Gene Ruler Express DNA ladder was loaded in a separate lane. As a negative control RNAs free water was used that replaced DNA. The gel was run at 90 volt for about 60 minutes. The gel was visualized under a UV transilluminator.

Next Generation Sequencing Analysis

Generated NGS library was sequenced using MiSeq sequencing (Illumina, San Diego, California) following the standard protocol. Bioinformatic analyses of the sequenced data were performed by using CLC Genomics workbench software version 20 (Qiagen, Denmark). Several thousand read matches of obtained sequences were found using local

realignment in CLC genomics workbench for each patient sample. For phylogenetic analysis, the full length BKV reference sequences were downloaded from GenBank and maximum to down to 10 % read match of the sequences, obtained in this study, were aligned with the reference sequence using Clustal W program implemented in MEGA version 6. Phylogenetic trees were constructed employing maximum likely hood method.

Results

In this study samples from four BKV positive patients who went through kidney transplantation were studied to identify viral co-infection and mutations. DNA samples from these four patients were purified and subjected to VP1 nested PCR in two rounds (**Figure 1**). In the first PCR (outer PCR), BKV VP1 gene of interest was amplified (**Figure 1**, upper panel). The amplicon of this PCR was then subjected to the second PCR (inner PCR) with two different strategies—one involved amplification using primers having index included (**Figure 1**, right bottom panel) whereas the other one was in the absence of index (**Figure 1**, left bottom panel). Samples from both preparations were then analyzed by NGS.

BK VP1 Nested PCR

Nested PCR was chosen to amplify gene of interest as well as to investigate whether nested PCR is required, or the regular PCR is sufficient to generate quality amplicons for further analysis. The extracted DNA from four different patient samples were subjected to BK VP1 nested PCR using two different sets of primers. The first round of PCR (outer PCR) using first set of primers aimed to yield 712 bp length amplicons and the second round of PCR using the other primers set was aimed to yield 327 bp length amplicons (the inner portion of VP1 gene).

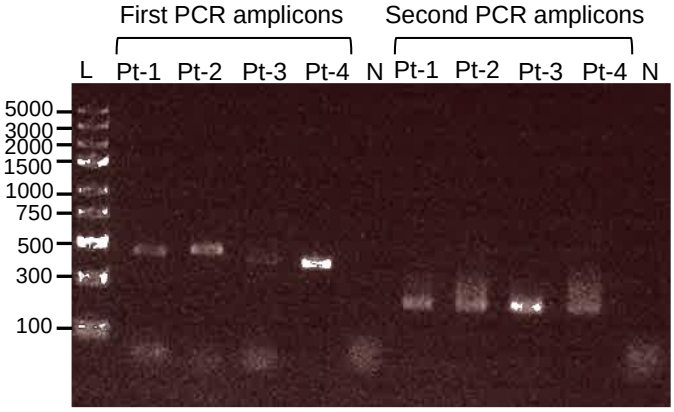

Figure 2: Agarose gel electrophoresis analysis showing the size of the amplicons obtained from the first and second round of BKV VP1 nested PCR. The designations L, Pt, and N refer to DNA ladder, patient, and negative control, respectively.

The size of the resulted amplicons was inspected using gel electrophoresis. The amplified products from the first round PCR appeared at about 500 bp (**Figure 2**, First PCR amplicons), and the products from second round PCR were appeared at about 250 bp (**Figure 2**, Second PCR amplicons). The appearance of DNA band on the gel is somewhat different than the desired size. Therefore, for further accurate measurement, Tape station analysis was performed. This analysis showed that amplicons from the fist round PCR were in the range of 754 bp–764 bp. Whereas amplicons from the second round PCR were in the range of 356 bp–360 bp. Representative images of Tape station analyses for patient 1 and 2 from first PCR and for patient 3 and 4 from second PCR are presented in **Figure 3**. These results are in line with the expectation. Furthermore, results from gel electrophoresis and tape station

measurement showed that both first and second rounds PCR produced quality amplicons and also indicate that amplicons from any of these two PCR rounds could be chosen for generation of NGS library.

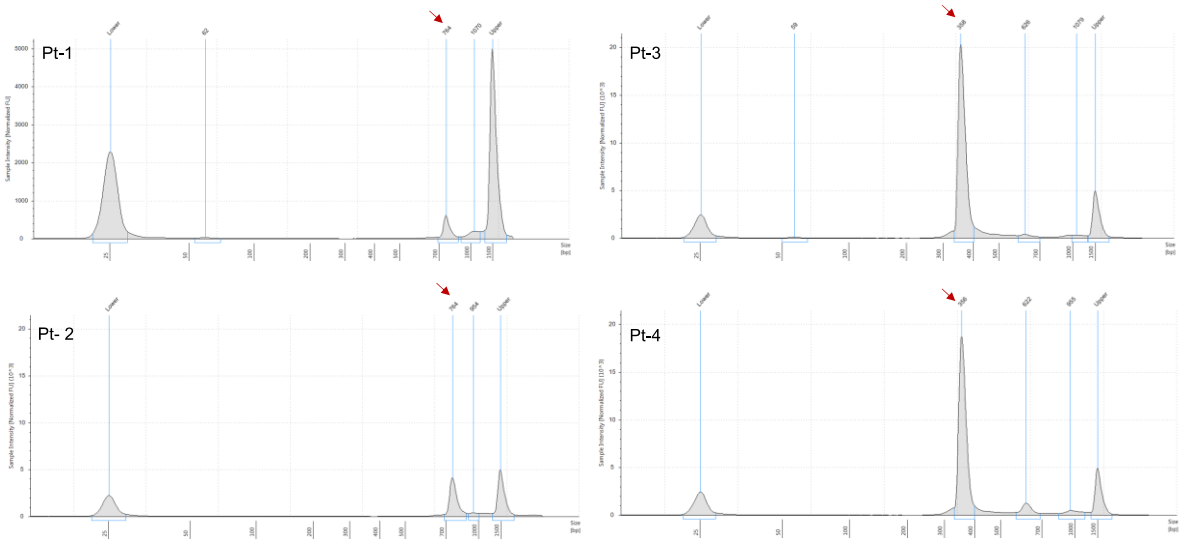

Figure 3: Representative images Tape station analysis showing the sizes of the amplicons resulted from BKV VP1 nested PCR. First round PCR amplicons for patients 1 and 2 were measured 764bp (peak indicated by arrowhead), and second round PCR amplicons for patients 3 and 4 were measured 358bp and 356bp, respectively (peak indicated by arrowhead). These four amplicons were selected for DNA library generation.

Next Generation Sequencing

We next performed NGS, which is the major interest of the study to know whether there are co-infections and mutations in patient sample analyzed. Prior to the NGS sequencing, two amplified PCR products from the first PCR round (patient 1 and 2) and two from second PCR round (patient 3 and 4) were selected for Nextera DNA library preparation. As because NGS sequencing success depends on accurate quantification of input DNA, the selected amplicons were subjected to Qubit DNA assay for accurate quantification as recommended by Nextera DNA library preparation protocol. The amount of DNA (ng/µl) measured in the four selected samples are presented in **Table 3.** A DNA library using these four samples was created and subsequently sequenced on a MiSeq sequencing (Illumina, Illumina, San Diego, California). No significance difference between the data obtained from the sequencing of first and second round PCR amplicons were noted, which further supports the observation made in previous section.

Table 3. Amount of DNA (ng/µl), in the selected sample for Nextera Library preparation.

First Round PCR (Longer DNA fragments)	
Sample	**DNA (ng/µl)**
Pt-1 (764bp)	1.59
Pt-2 (764bp)	9.72
Second Round PCR (Shorter DNA fragments)	
Sample	**DNA (ng/µl)**
Pt-3 (358bp)	42.0
Pt-4 (356bp)	42.6

Phylogenetic Analysis of BKV VP1 Sequences

Since one of the aims of this project was to identify the presence of co-infection, we performed phylogenetic analysis to investigate whether any of the patient studied were co-infected with more the one BKV genotype or subtype. The local realignment tools in CLC genomic workbench produced thousands of reads matches. The maximum to down to 10 % read match of each patient sample was compared with full length BKV reference sequences downloaded from GenBank®. The read matches of each of the patient was found to be best aligned with one BKV genotype or subtype. A phylogenetic tree (**Figure 4**) was constructed by aligning maximum read match of four patients against reference gene.

Based on VP1 sequence read match, the patient samples, denoted as Pt-1 to Pt-4, were grouped under genotype BKV I and IV (**Figure 4**). VP1 sequence read match from Pt-1, Pt-2, and Pt-4 were best aligned and grouped into BK Ib_2 subtype and sequence read match from Pt-3 was grouped under Bk IVb_2. According to phylogenetic analysis, no more than one BKV genotype was identified in any of the patient sample analysed.

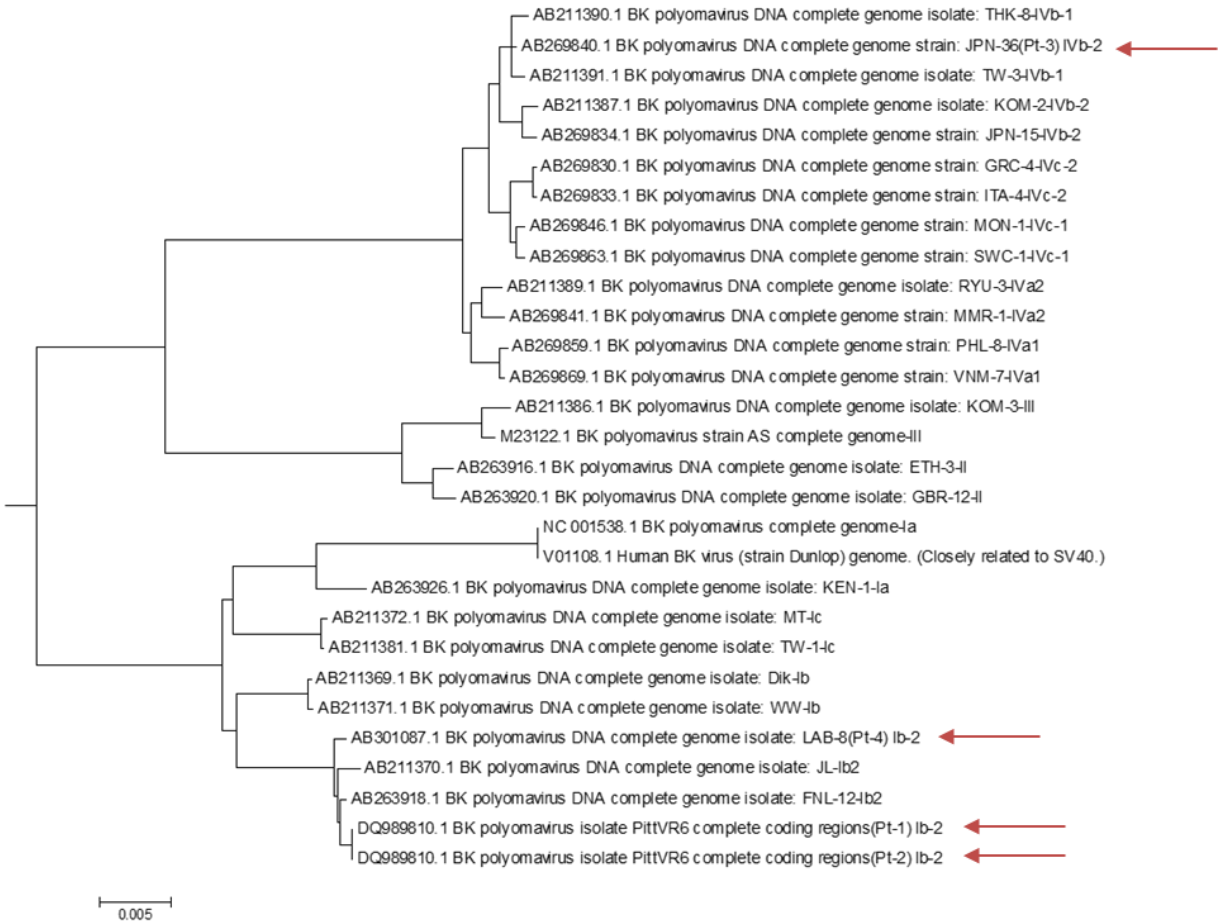

Figure 4: The maximum likelihood phylogenetic tree of BKV genome. Phylogenetic tree was based on maximum read match of 4 BKV positive sample (Pt-1, Pt-2, Pt-3, Pt-4) from four kidney transplant recipient and 24 BKV reference genome, annotated with GenBank accession numbers. Samples are indicated with patient number within parenthesis (for example, Pt-3 IVb2) and also marked with red arrowhead.

Mutation Analysis of BKV VP1 Sequences

In attempt to detect if the virus which infected the patients have had acquired any changes to escape immune system, sequence obtained from NGS were analysed in CLC genomics workbench. Extracted consensus sequences were aligned and assembled against reference strain using Clustal W in MEGA (version 6) with default parameters to compare and detect the presence of mutation and other sequence variants. A selection of the alignments is shown in **Figure 5**. Mutations were detected in multiple reads within the VP1 region in all tested samples, except one sample from patient three (Pt-3), and some of these mutations are results in amino acids changes in the respective positions, whereas others are silent mutations (**Table 4**). In patient one, a silent and a point mutation were detected at positions 6 and 7, respectively. The lysine residue at position 7 was found to be mutated to arginine. Two of the four mutations detected in patient two (at position 27 and 125) were silent, and the other two were point mutations, which were detected in different reads the at same position i.e., 84. The two possible changes of glutamic acid at this position are aspartic acid and lysine. Remarkably, the greatest number of mutations i.e., five were detected in patient four. Two of these, at position 27 and 44, were found to be silent mutations. The other tree glutamic acid, lysine, and asparagine at positions 84, 86, and 102, respectively were mutated to aspartic acid, arginine, and serine.

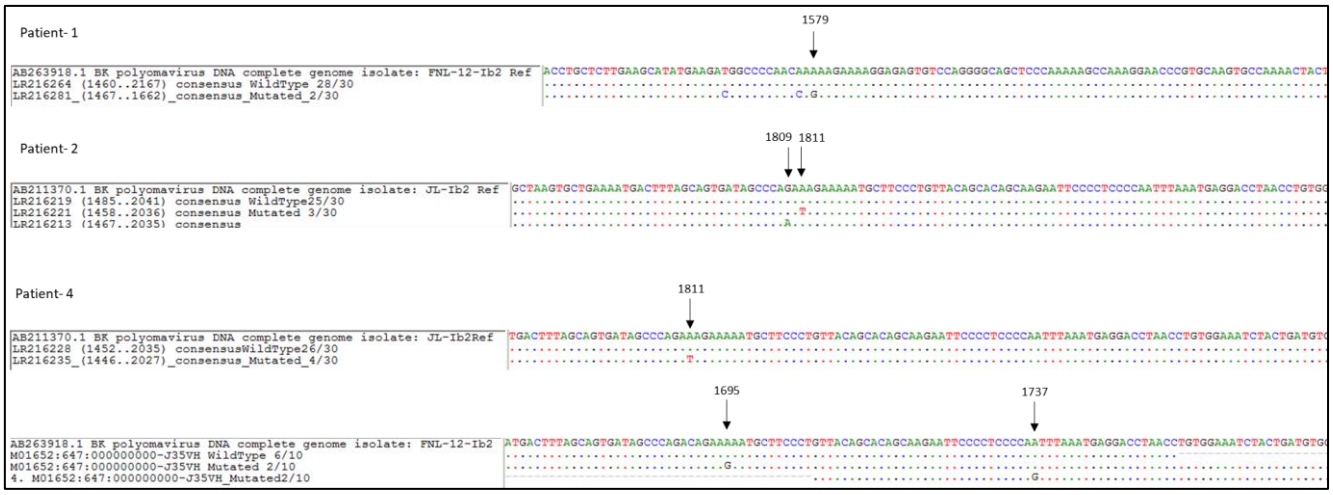

Figure 5: Figure showing the representative alignments of wildtype and mutated consensus sequence against reference strain. Alignment for six point mutations listed in **Table 4** are shown. Positions of mutated nucleotide are labelled.

Table 4. Representation of the summery of mutations detected in VP1 region by deep sequencing of plasma sample from four kidney transplant recipients. The colour code 'cyan' represents changed from and 'magenta' changed to. The position of the amino acid is mentioned below the arrow in amino acid column.

Patient	Number of mutation	Mutations trinucleotide	Amino acid change	No. of reads
Pt-1	2	ACA → ACC	Thr → Thr 6	4
		AAA → AGA	Lys → Arg 7	2
Pt-2	4	CTA → CTG	Leu → Leu 27	3
		GAA → GAT	Glu → Asp 84	3
		GAA → AAA	Glu → Lys 84	2
		AGC → AGT	Ser → Ser 125	18
Pt-3	–	–	–	–
Pt-4	5	CTA → CTG	Leu → Leu 27	4
		CTA → TTA	Leu → Leu 44	4
		GAA → GAT	Glu → Asp 84	4
		AAA → AGA	Lys → Arg 86	2
		AAT → AGT	Asn → Ser 102	2

MiSeq Amplicon Sequencing

To set up a detection procedure in parallel with NGS followed by XT DNA library generation, we prepared sample using Miseq duel indexing amplicon preparation. However, the sequencing of samples from this preparation was not performed with in the scope of this study.

The MiSeq amplicon sequencing sample preparation procedure generates barcoded amplicon due to the addition of index primer, which is ready to be sequenced right after nested PCR thereby omit the necessity of a tagmented library preparation. We utilized this advantage and prepared sample using MiSeq duel indexing library preparation. The VP1 region of genomic DNA was amplified using a VP1 nested PCR with two different set of primers. The first amplification round was carried out to amplify 712 bp length amplicons followed by a second round tagged PCR carried out with a set of primers attached to duel indexes to barcode the amplicons in order to be able to recognize sample after sequencing multiple samples at the

same time. The size of the amplicons was measured using gel electrophoresis (**Figure 6**). DNA bands for all patient samples were appeared around 350 bp. As in the previous preparation, amplicons from this preparation were also analysed using Tape station. The measured size of the amplicons was in the range of 547 bp to 572 bp (**Figure 7**).

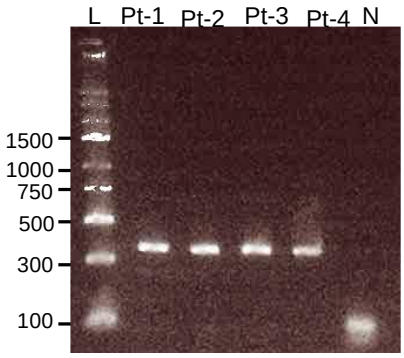

Figure 6: Agarose gel electrophoresis analysis showing the size of amplicons obtained after a second round tagged PCR with duel indexed primers. Samples from patient 1 was tagged with indexes iF11v3BK and iR11v3BK, from patient 2 was tagged with indexes iF12v3BK and iR12v3BK4, from patient 3 was tagged with indexes iF13v3BK and iR13v3BK, and from patient 4 was tagged with indexes iF15v3BK and iR15v3BK. The designations L, Pt, and N refer to DNA ladder, patient, and negative control, respectively.

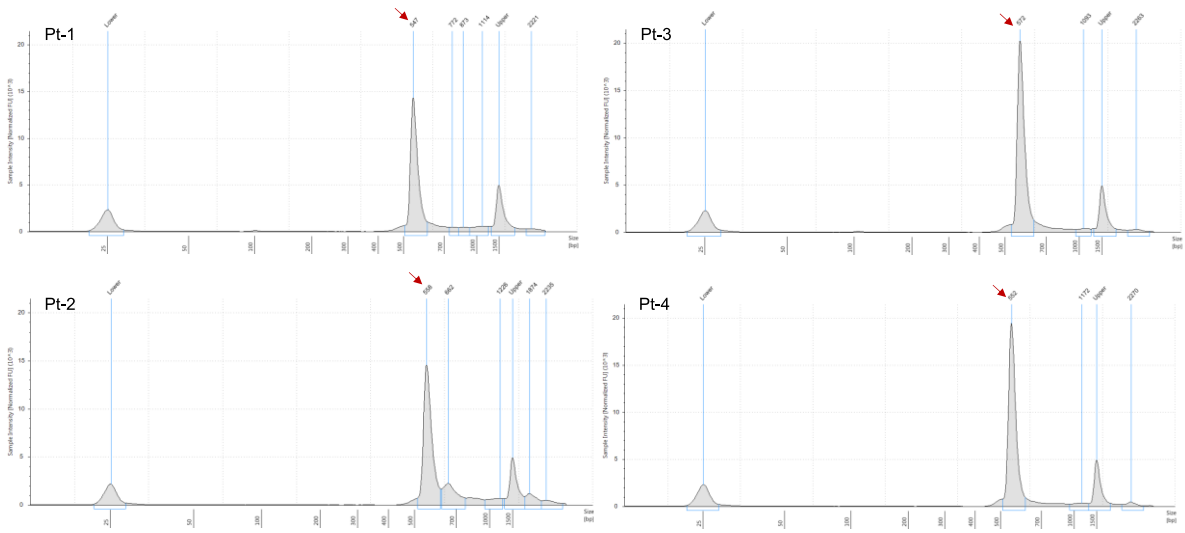

Figure 7: Representative images of Tape station analysis showing the more accurate size of the indexed amplicons resulted from tagged PCR. The amplicon size for patients 1, 2, 3, and 4 was measured 547 bp, 558 bp, 572 bp, and 552 bp, respectively. The arrowhead indicates the corresponding peak.

Discussion

BKV infection in renal transplant recipients is a crucial challenge globally. The reactivation of BKV is triggered by the immune suppression can potentially lead to the development of PVAN, which a major post-transplant complication causing allograft dysfunction (40). However, in addition to immune suppression, and other general risk factors such as age, HLA mismatch or diabetes, the viral risk factors are believed to play critical roles in the development of PVAN (25,41).

Herein we analysed plasma samples from four renal transplant recipients by using NGS to identify potential co-infections (virus from donor organ) and mutations in the VP1 gene of BKV. The DNA extracted from patient samples was amplified and sequenced using NGS. We chose to work with NGS since this method is superior to the standard Sanger sequencing method in terms of high quantity data outcomes, cost-effectiveness, and reduction in analyses time (38,39).

We performed phylogenetic analysis to investigate whether any of these four patients have more than one subtype or genotype of BKV, which is a hint for possible co-infections with donor virus. We have detected genotype I (subtype Ib2) in three out of the four patients and genotype IV (subtype IVb2) in one patient. None of the patients were detected with more the one BKV genotype or subtype. Our result is not in agreement with previous studies. Using a similar strategy as ours, two different BKV genotype/subtype (BKV genotype I [subtype Ib_2] and IV) were identified in urine sample of hematopoietic stem cell transplant recipient (32). Also, a neutralization serology assay of patient and matched kidney donor sera showed that donor sera effectively neutralized two different BKV genotypes, indicating that the donor was co-infected with two BKV genotypes (31). Thus, patients in this study might be infected with the same BKV genotype/subtype as the donor. Alternatively, the lack of co-infection might be the result of the small cohort used in this study and a larger sample size might reveal different findings.

Next, we have identified multiple point mutations and some of these mutations result in changes in amino acid residues in the corresponding position of VP1 gene. We have analysed 30 reads for each patient, and mutation which was found in more than one read was considered. In total, six missenses and five silent mutations were found in different samples tested. Our results resonate well with earlier studies (41,42) that reported at least two mutation in the same genomic region. However, the clinical significance of these identification is unknown.

In this study, three out of six mutations that alter codon to a different amino acids were identified within the subtyping region of the BC loop and rest of the mutations were randomly distributed, which is not uncommon (43). Although we have studied only limited numbers of reads from each patient, multiple random point mutations were detected in most of the patient. Moreover, mutation rate in DNA viruses is known to be lower compared to RNA viruses. Furthermore, we believe that the extensive analysis (which is out of the scope of the present study) of the data obtained from NGS might provide more even promising outcome to understand infection panorama.

In conclusion, we have presented a methodology where PCR amplification was coupled to high-throughput NGS analyses for the detection of viral co-infection and minor variants in clinical samples. We used BK polyomavirus as a model system. Results generated in this report demonstrated that this methodological approach has potential. However, the approach should be tested in a large sample cohort for further investigations of its usefulness in clinical practice as well as validation.

Acknowledgement

I would like to express my gratitude to Dr. Christina Öhrmalm and Dr. Anders Bergqvist for the opportunity of this master thesis project at the department of Clinical Microbiology, Uppsala University Hospital. Moreover, I would like to thank them for their ongoing support throughout the project. I would like to extend my gratitude to Dr. Lucia Cavelier and Britt E Inger Jonsson for all the technical support with NGS and effective collaboration.

Reference:

1. Vanichanan J, Udomkarnjananun S, Avihingsanon Y, Jutivorakool K. Common viral infections in kidney transplant recipients. Kidney Res Clin Pract. 2018 Dec 31;37(4):323–37.

2. Gardner SD, Field AM, Coleman DV, Hulme B. New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet. (i:1253–1257.).

3. Gardner SD. Prevalence in England of antibody to human polyomavirus (B.k.). Br Med J. (1):77–78.

4. Discovery and epidemiology of the human polyomaviruses BK virus (BKV) and JC virus (JCV). Adv Exp Med Bio. (577):19 – 45.

5. Doerries K. Human polyomavirus JC and BK persistent infection. Adv Exp Med Biol. (577):102–116.

6. Murray, R.; Rosenthal, K.S.; Kobayashi,G.S. Medical microbiology. Vol. 4ed.2002. Mosby,inc;

7. Yogo Y, Sugimoto C, Zhong S, Homma Y. Evolution of the BK polyomavirus: epidemiological, anthropological and clinical implications. Rev Med Virol. 2009 Jul;19(4):185–99.

8. DeCaprio JA, Garcea RL. A cornucopia of human polyomaviruses. Nat Rev Microbiol. 2013 Apr;11(4):264–76.

9. Bohl DL, Brennan DC. BK Virus Nephropathy and Kidney Transplantation. CJASN. 2007 Jul;2(Supplement 1):S36–46.

10. Hurdiss DL, Morgan EL, Thompson RF, Prescott EL, Panou MM, Macdonald A, et al. New Structural Insights into the Genome and Minor Capsid Proteins of BK Polyomavirus using Cryo-Electron Microscopy. Structure. 2016 Apr;24(4):528–36.

11. Tremolada S, Delbue S, Castagnoli L, Allegrini S, Miglio U, Boldorini R, et al. Mutations in the external loops of BK virus VP1 and urine viral load in renal transplant recipients. J Cell Physiol. 2010 Jan;222(1):195–9.

12. Chatterjee M, Weyandt TB, Frisque RJ. Identification of archetype and rearranged forms of BK virus in leukocytes from healthy individuals. J Med Viro. (60):353–362.

13. Neu U, Stehle T, Atwood WJ. The Polyomaviridae: Contributions of virus structure to our understanding of virus receptors and infectious entry. Virology. 2009 Feb;384(2):389–99.

14. Byun H, Gou Y, Zook A, Lozano MM, Dudley JP. ERAD and how viruses exploit it. Front Microbiol [Internet]. 2014 Jul 3 [cited 2020 Apr 6];5. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2014.00330/abstract

15. Eash S, Querbes W, Atwood WJ. Infection of Vero Cells by BK Virus Is Dependent on Caveolae. Journal of Virology. 2004 Nov 1;78(21):11583–90.

16. Zheng HY, Nishimoto Y, Chen Q, Hasegawa M, Zhong S, Ikegaya H, Ohno N, Sugimoto C, Takasaka T, Kitamura T, Yogo Y. Relationships between BK virus lineages and human populations. Microbes Infect. (9):204–13.

17. Ikegaya H, Saukko PJ, Tertti R, Metsarinne KP, Carr MJ, Crowley B,, Sakurada K, Zheng HY, Kitamura T, Yogo Y. . Identification of a genomic subgroup of BK polyomavirus spread in European populations. J Gen Virol. (87):3201–8.

18. Zhong S, Randhawa PS, Ikegaya H, et a. Distribution patterns of BK polyomavirus (BKV) subtypes and subgroups in American, European and Asian populations suggest co-migration of BKV and the human race. J Gen Virol. (90):144–52.

19. Ambalathingal GR, Francis RS, Smyth MJ, Smith C, Khanna R. BK Polyomavirus: Clinical Aspects, Immune Regulation, and Emerging Therapies. Clin Microbiol Rev. 2017 Apr;30(2):503–28.

20. Goudsmit J, Wertheim-van Dillen P, van Strien A, van der Noordaa J. The role of BK virus in acute respiratory tract disease and the presence of BKV DNA in tonsils. . J Med Viro. (10):91–9.

21. Heritage J, Chesters PM, McCance DJ. The persistence of papovavirus BK DNA sequences in normal human renal tissue. J Med Virol. (8):143–50.

22. Bressollette-Bodin C, Coste-Burel M, Hourmant M, Sebille V, Andre-Garnier E, Imbert-Marcille BM. A Prospective Longitudinal Study of BK Virus Infection in 104 Renal Transplant Recipients. Am J Transplant. 2005 Aug;5(8):1926–33.

23. Broekema NM, Imperiale MJ. miRNA regulation of BK polyomavirus replication during early infection. Proceedings of the National Academy of Sciences. 2013 May 14;110(20):8200–5.

24. Tian YC, Li YJ, Chen HC, Wu HH, Weng CH, Chen YC, Lee CC, Chang MY, Hsu HH, Yen TH, Hung CC, Yang CW. Polyomavirus BK-encoded microRNA suppresses autoregulation of viral replication. Biochem BIophys Res Commun. (447):543–9.

25. Boothpur R, Brennan DC. Human polyoma viruses and disease with emphasis on clinical BK and JC. Journal of Clinical Virology. 2010 Apr;47(4):306–12.

26. Reploeg MD, Storch GA, Clifford DB. BK Virus: A Clinical Review. CLIN INFECT DIS. 2001 Jul 15;33(2):191–202.

27. Hirsch HH, Randhawa P, the AST Infectious Diseases Community of Practice. BK Polyomavirus in Solid Organ Transplantation: BK Polyomavirus in Solid Organ Transplantation. American Journal of Transplantation. 2013 Mar;13(s4):179–88.

28. Bennett SM, Broekema NM, Imperiale MJ. BK polyomavirus: emerging pathogen. Microbes and Infection. 2012 Aug;14(9):672–83.

29. Nickeleit V, Hirsch HH, Binet IF, Gudat F, Prince O, Dalquen P, et al. Polyomavirus Infection of Renal Allograft Recipients: From Latent Infection to Manifest Disease. J Am Soc Nephrol. :10.

30. Solis M, Velay A, Porcher R, Domingo-Calap P, Soulier E, Joly M, Meddeb, M, Kack-Kack W, Moulin B, Bahram S, Stoll-Keller F, Barth H, Caillard S,, Fafi-Kremer S. Neutralizing antibody-mediated response and risk of BK virus-associated nephropathy. J Am Soc Nephrol. (29):326 –334.

31. Peretti A, Geoghegan EM, Pastrana DV, Smola S, Feld P, Sauter M, et al. Characterization of BK Polyomaviruses from Kidney Transplant Recipients Suggests a Role for APOBEC3 in Driving In-Host Virus Evolution. Cell Host & Microbe. 2018 May;23(5):628-635.e7.

32. Van TD, Dwyer DE, Jeoffreys N, Kok J, Nankivell BJ, Rockett RJ. Deep Sequencing of Urine Specimens Detects Two BK Polyomavirus Genotypes in a Hematopoietic Stem Cell Transplant Recipient. Gill SR, editor. Microbiol Resour Announc. 2020 Jan 30;9(5):e01461-19, /mra/9/5/MRA.01461-19.atom.

33. Wiseman AC. Polyomavirus nephropathy: a current perspective and clinical considerations. Am J Kidney Dis. 54(1):131–142.

34. Wunderink HF, de Brouwer CS, van der Meijden E, Pastrana DV, Kroes ACM, Buck CB, et al. Development and evaluation of a BK polyomavirus serotyping assay using Luminex technology. Journal of Clinical Virology. 2019 Jan;110:22–8.

35. Comoli P, Cioni M, Basso S, Gagliardone C, Potenza L, Verrina E, et al. Immunity to Polyomavirus BK Infection: Immune Monitoring to Regulate the Balance between Risk of BKV Nephropathy and Induction of Alloimmunity. Clinical and Developmental Immunology. 2013;2013:1–6.

36. Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, Buck CB. BK polyomavirus genotypes represent distinct serotypes with distinct entry tropism. J Virol. 87:10105–10113.

37. Randhawa P, Khaleel-Ur-RehmanK, Swalsky PA, Vats A, Scantlebury, V, Shapiro R, Finkelstein S. DNA sequencing of viral capsid protein VP-1 region in patients with BKvirus interstitial nephritis. Transplantation. 73:1090–1094.

38. Wong L-JC. Challenges of bringing next generation sequencing technologies to clinical molecular diagnostic laboratories. Neurotherapeutics. 2013 Apr;10(2):262–72.

39. Dziubianau M, Hecht J, Kuchenbecker L, Sattler A, Stervbo U, Rödelsperger C, et al. TCR Repertoire Analysis by Next Generation Sequencing Allows Complex Differential Diagnosis of T Cell-Related Pathology: NGS Allows Complex Differential Diagnosis. American Journal of Transplantation. 2013 Nov;13(11):2842–54.

40. Hirsch HH, Brennan DC, Drachenberg CB, Ginevri F, Gordon J, Limaye AP, et al. Polyomavirus-Associated Nephropathy in Renal Transplantation: Interdisciplinary Analyses and Recommendations: Transplantation. 2005 May;79(10):1277–86.

41. Boldorini R, Allegrini S, Miglio U, Paganotti A, Veggiani C, Mischitelli M, et al. Genomic mutations of viral protein 1 and BK virus nephropathy in kidney transplant recipients. J Med Virol. 2009 Aug;81(8):1385–93.

42. Jin L, Gibson PE, Booth JC, Clewley JP. Genomic typing of BK virus in clinical specimens by direct sequencing of polymerase chain reaction products. J Med Virol. 1993;41:11–7.

43. Liimatainen H, Weseslindtner L, Strassl R, Aberle SW, Bond G, Auvinen E. Next-generation sequencing shows marked rearrangements of BK polyomavirus that favor but are not required for polyomavirus-associated nephropathy. Journal of Clinical Virology. 2020 Jan;122:104215.