THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

DETECTION AND CHARACTERIZATION OF VIRAL cDNAs AMPLIFIED FROM CELLS EXPRESSING THE JC VIRUS EARLY CODING SEQUENCES

ERIN STOCKWELL BARNES
Spring 2011

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Richard Frisque, Ph.D.
Professor of Molecular Virology
Thesis Supervisor

Craig Cameron, Ph.D.
Paul Berg Professor of Biochemistry and Molecular Biology
Honors Adviser

Wendy Hanna-Rose, Ph. D.
Associate Department Head for Undergraduate Studies
Department of Biochemistry and Molecular Biology

* Signatures are on file in the Schreyer Honors College.

Abstract

The early region of human polyomavirus JC virus (JCV) produces five proteins, large T antigen (TAg), small t antigen (tAg), and T´$_{135}$, T´$_{136}$, and T´$_{165}$, encoded by transcripts alternatively spliced from one pre-mRNA. These multifunctional regulatory proteins mediate viral DNA replication and contribute to the oncogenic potential of the virus. Initially proposed to be proteolytic cleavage products of JCV TAg, the three T´ proteins have been shown to be authentic viral phosphoproteins that participate in JCV's viral activities. T´ transcripts are generated by the removal of two introns from unspliced JCV early pre-mRNA; the first intron also being removed in the TAg transcript. Since the splice variants share their amino termini with TAg, all three T´ proteins contain the Rb binding motif LxCxE and the Hsc70 binding J domain, and thus exhibit some notable TAg functions such as stimulation of cell proliferation. The T´ cDNAs were first observed on polyacrylamide gels in RT-PCR experiments, in which RNA was extracted from JCV-infected cells, reverse transcribed into cDNA, and amplified by PCR. Five bands representing cDNAs encoding the JCV early T proteins were visualized on a polyacrylamide gel, two slower migrating "doublet" bands, and three faster migrating T´ bands 17-23 kD in size. Based upon the primers used, the TAg and tAg cDNAs are expected to be the same size and therefore form a single band positioned above the three T´ bands. The appearance of two bands has yet to be explained in an experimental setting, so to unravel this mystery, I have attempted to clone and identify the origin of these sequences within the cDNA doublet bands. Through a number of experimental approaches, our results have shown the repeated detection of the T´$_{136}$ splice pattern within the cDNAs isolated from the doublet band. We believe that a small amount of T´$_{136}$ cDNA remains "trapped" in the larger TAg/tAg cDNAs, and is preferentially selected for detection by our experimental methods. Further investigation is needed to conclude that TAg/tAg is the predominanat species in one of the doublet bands. Additional analysis will also be required to determine the origin of the non-TAg/tAg cDNA doublet band.

Table of Contents

Introduction

JC virus (JCV) is a human polyomavirus that establishes persistent infections in approximately 50 percent of the human population worldwide (Kean el al., 2009). Though the virus usually persists in the body throughout a person's lifetime, infection is usually asymptomatic and only causes symptoms in severely immunodeficient patients. AIDS patients and transplant recipients are the most common groups to develop the JCV-mediated fatal demyelinating disease of the central nervous system (CNS) called progressive multifocal leukoencephalopathy (PML). While PML was once considered a rare disease, it is now estimated that 5% of AIDS patients succumb to the neurological disease. Recently PML has occurred in individuals taking the drug Tysabri as treatment for Crohn's disease and multiple sclerosis (Hutchinson, 2007). Tysabri (natalizumab) is believed to work by binding receptors on B and T cells to prevent the migration of these immune cells out of the bloodstream and into the tissues. Because the drug inhibits lymphocytes from passing through the blood-brain barrier, researchers speculate that immune surveillance is reduced. If JCV is already present and latent in the brain as proposed by several laboratories (Elsner and Dorries, 1992; White *et al*, 1992; Mori *et al*, 1991), then reactivation of the virus might occur, leading to PML. PML results from the lytic destruction of oligodendrocytes, the myelin producing glial cells in the CNS. The demyelination of white matter in the brain genernally results in death within six to nine months of the onset of neurological symptoms (Walker *et al*, 1973).

The primary site of JCV DNA infection is unknown, but it is clear that the kidney is a major secondary site of infection. Viral particles are shed in the urine, but controversy remains regarding how and when JCV reaches the brain to cause disease. JCV is also known to have oncogenic potential, reflective of the group name *polyoma* meaning "many tumors." Cell culture and animal experiments have demonstrated JCV's oncogenic ability, and JCV has been found in

several human cancers, including brain, lung, and colon tumors (Frisque *et al*, 2006). JC virus' potential as a human tumor virus and as an emerging opportunistic pathogen that causes PML make it an important and urgent focus of research.

JCV is a member of the family *Polyomaviridae*, characterized by small, non-enveloped DNA viruses containing early and late genes regulated by a non-coding control region. The early regions of polyomaviruses encode the small and large T-antigens, proteins that promote viral DNA replication and initiate oncogenic transformation. The late coding regions of these viruses specify the regulatory agnoprotein and the capsid proteins VP1, VP2, and VP3. JCV is closely related to the human polyomavirus BK virus (BKV) and the monkey polyomavirus Simian Virus 40 (SV40), sharing approximately 75% or 69% overall sequence homology, respectively. Despite their close genetic similarity, all three viruses have distinct biological properties including different host ranges, tissue tropisms, and pathogenic and oncogenic potentials (Frisque *et al*, 1984). Compared to SV40, JCV exhibits limited infectivity of cells in culture, reproducing efficiently only in primary human fetal glial (PHFG) cells. Furthermore, JCV transforms cells in culture much less efficiently than BKV and SV40. The biological variation of these closely related viruses is, in part, accounted for by differences within the viral early coding regions (Chuke et al, 1986; Bollag et al, 1989; Haggerty et al, 1989).

JCV and BKV were discovered in 1971 and remained the only known human polyomavriruses until 2007. In the last four years, sequences have been cloned from human tissues that represent seven new polyomviruses. These include KI virus (KIV), WU virus (WUV), Merkel cell virus (MCV), HPyV6, HPyV7, *Trichodysplasia spinulosa* virus (TSV), and HPyV9 (Allander *et al*, 2007; Gaynor *et al*, 2007; Feng *et al*, 2008; Schowalter *et al*, 2010; van der Meijden *et al* 2010; Scuda *et al*, 2011). MCV has been closely linked to Merkel cell carcinoma (MCC), a rare but aggressive skin cancer. Several studies have confirmed that approximately 80%

of Merkel cell carcinomas test positive for the presence of MCV DNA (Busam *et al*, 2009; Duncavage *et al*, 2009; Kassem *et al*, 2008). Tumor cells derived from MCV-positive MCC cases have been shown to express a truncated large T antigen (Houben *et al*, 2010).

JC virus was first isolated from autopsied brain tissue obtained from a patient with PML in Madison, Wisconsin, hence this prototype strain's denotation as Mad-1 (Padgett, et al, 1971). Sequenced in 1984, the JCV genome is a circular double-stranded DNA that consists of 5130 base pairs (See Figure 1) (Frisque *et al*, 1984). The noncoding regulatory region contains cis-acting elements, including promoter/enhancer transcription elements and the origin of viral DNA replication. The early region produces five proteins; large T antigen (TAg), small t antigen (tAg), T'_{135}, T'_{136}, and T'_{165}, that are translated from transcripts alternatively spliced from the same early pre-mRNA. The early and late genes are transcribed in opposite directions, and are thus encoded by opposite strands of DNA (Trowbridge and Frisque, 1995).

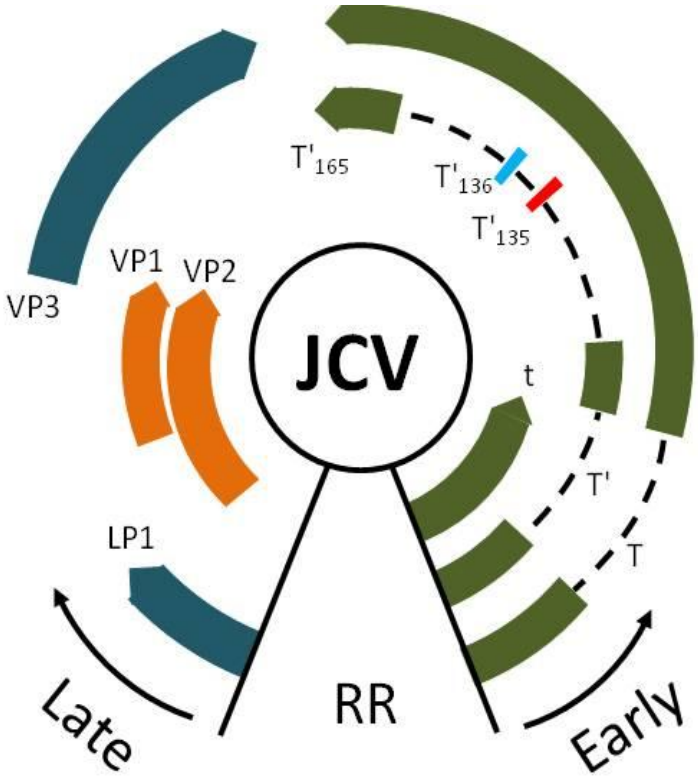

Figure 1. Map of JC Virus Genome. The diagram above depicts a representation of the viral genome and encoded proteins of JCV. The complete JCV sequence is 5130 nucleotides in length and, in total, codes for 9 proteins. The oncogenic T family of early proteins consists of large T Antigen (TAg), small t antigen (tAg), T´135, T´136, and T´165. Late proteins include the regulatory agnoprotein (LP1), and viral capsid proteins VP1,VP2, and VP3. The regulatory region (RR) is the non-coding sequence consisting of the origin of replication and the promoter/enhancer elements for transcription (Trowbridge and Frisque, 1995) (modified from Frisque, 2001).

There are three known types of sequence variation within the JCV genome. The first kind of variation involves single nucleotide polymorphisms (SNPs) that occur within the genome's coding region. Evidence of coding region variability amongst the fifteen different genotypes and subtypes suggests long term viral evolution in the human host. Specific genotypes, traced to specific geographic/ethnic populations, have been identified that exhibit DNA sequence deviation of 1-3 percent amongst different variants (Cui *et al*, 2004; Jeong *et al*, 2004). Coding region variability has been correlated with differences in pathogenicity and oncogenicity

(Frisque, 2003). A second type of variation involves sequences within the promoter/enhancer elements, and is represented by the archetype and rearranged variants. JCV archetype is the form that persists in the kidneys, is excreted in the urine, and infects new susceptible individuals. Archetype JCV does not grow in PHFG cells in culture and it is not associated with diseased tissues. All isolates of archetype JCV derived from humans essentially have identical promoter/enhancer sequences. In contrast, there are multiple rearranged forms of JCV that all differ from archetype by a series of deletions and duplications within the promoter/enhancer. The rearrangements appear to occur within the host following infection with the archetype form. Rearranged forms grow in PHFG cells and are found in PML brain tissue (Frisque and White, 1992). A third JCV variant has been described only recently. In this case, the VP1 gene of the JCV genomes isolated from the brain and kidneys of a single PML patient differs by only a few SNPs. Alterations to the VP1 gene in the brain isolates are thought to lead to changes in cell receptor interactions and immune recognition (Sunyav *et al*, 2009).

TAg is the major polyomavirus regulatory protein, and it is highly conserved amongst the members of this virus family (Johnson, 2010). TAg mediates viral DNA replication in permissive cells and induces oncogenic alterations in non-permissive cells (Frisque and White III, 1992). Studies show that TAg promotes host cell cycle progression in order to induce the expression of replication factors required during S phase. In permissive human cells, this activity supports viral DNA replication, but in non-permissive cells, this activity may lead to uncontrolled cellular proliferation and oncogenic transformation (Bollag *et al*, 2006). These and numerous other activities of TAg have been mapped to a large number of functional domains within the protein. These domains facilitate interactions with multiple cellular proteins that include the Rb pocket proteins, the tumor suppressor p53, the molecular chaperone hsc70, and the Wnt signaling molecule β-catenin (Brodsky and Pipas, 1998; Bollag et al, 2000; Mendoza and Frisque, 2011).

A large body of work has focused upon the interactions of polyomavirus T protein with members of the Rb (Retinoblastoma) family, a group of tumor suppressor proteins that regulate cell cycle progression. Unphosphorylated Rb proteins are active in the G_0/G_1 phase, and prevent cellular progression into S phase by forming Rb-E2F complexes that inhibit E2F-dependent transcription. Rb inactivation through cyclin-dependent kinase-mediated phosphorylation allows for the release of E2F factors that promote transcription of genes required for cellular DNA synthesis. The amino terminus of the JCV TAg contains an LxCxE motif, which binds to the B pocket of Rb proteins. A second amino terminal sequence in the JCV TAg, called the J domain, binds to the cellular molecular chaperone hsc70 and activates its intrinsic ATPase activity, thereby causing the release of E2F from the Rb-E2F complex (Sheng *et al,* 1997). In this way, the JCV TAg overcomes the normal regulated process of Rb phosphorylation to effect unscheduled S phase progression. The oncogenic behavior of a polyomavirus TAg is due, in part, to its ability to inactivate this key tumor suppressor protein and promote uncontrolled cell proliferation.

The polyomavirus TAg, under some conditions, is the only viral protein required for viral DNA replication and cellular transformation. However, tAg may also contribute to these processes. As with TAg, the tAg is conserved among the primate polyomavriruses. Its unique carboxy-terminal sequences bind to a critical cellular phosphatase, PP2A. The regulatory B subunit of PP2A is replaced by tAg, leading to multiple downstream effects involving activation of the MAP kinase pathway and stabilization of transcription factors (Cho *et al,* 2007). In some ways, tAg acts like a growth factor by promoting cellular proliferation. Recent studies indicate that the JCV t protein exhibits unique functions not observed in most other polyomaviruses, such as being required for efficient viral DNA replication in cell culture. While this effect may involve the PP2A interaction, a second property of the JCV tAg, the ability to bind Rb proteins, may play a role. This property was only demonstrated recently, following the observation that JCV tAg,

unlike other tAgs, has two LxCxE motifs in its carboxy-terminus. To date, no other polyomavirus tAg has been demonstrated to bind Rb proteins (Bollag *et al,* 2010).

The three T´ proteins were initially proposed to be proteolytic cleavage products of JCV TAg (Bollag *et al*, 1989; Haggerty *et al*, 1989). Several years later, it was reported that these 17-23kD phosphoproteins, first detected on polyacrylamide gels, were, in fact, three authentic viral proteins encoded by alternatively-spliced early mRNAs. T'_{135}, T'_{136}, and T'_{165}, named after their respective amino acid lengths, are produced by the same alternative splicing mechanism that yields TAg and tAg (see Figure 2). T´ transcripts are generated by the removal of two introns from unspliced JCV early pre-mRNA; the first intron is also removed in the TAg transcript. The second intron in T´ mRNAs is removed by utilizing a shared 5´ donor splice site at nucleotide #4274 and three different 3´ acceptor sites at # 2918, 2777, and 2704 for T'_{135}, T'_{136}, and T'_{165}, respectively (Trowbridge and Frisque, 1995). Produced are three "truncated" T proteins that are 135, 136, and 165 amino acids in length and which share their initial 132 amino acids with TAg. T'_{135} and T'_{136} have unique carboxy termini, whereas T'_{165}, which is translated in the same reading frame as TAg, shares its carboxy-terminal 33 amino acids with TAg. Because their amino termini are shared with TAg, all three T´ proteins contain the LxCxE and J domains, known binding motifs of Rb family proteins and Hsc70, respectively. This observation suggests that the T´ splice variants would exhibit some TAg functions, such as stimulation of cell proliferation. This prediction is supported by studies showing that T´ proteins differentially bind to pRb, p107, and p130 *in vitro* and *in vivo* (Bollag *et al,* 2000, Bollag *et al*, 2006), and exhibit immortalizing and transforming behaviors (Bollag *et al,* 2010).

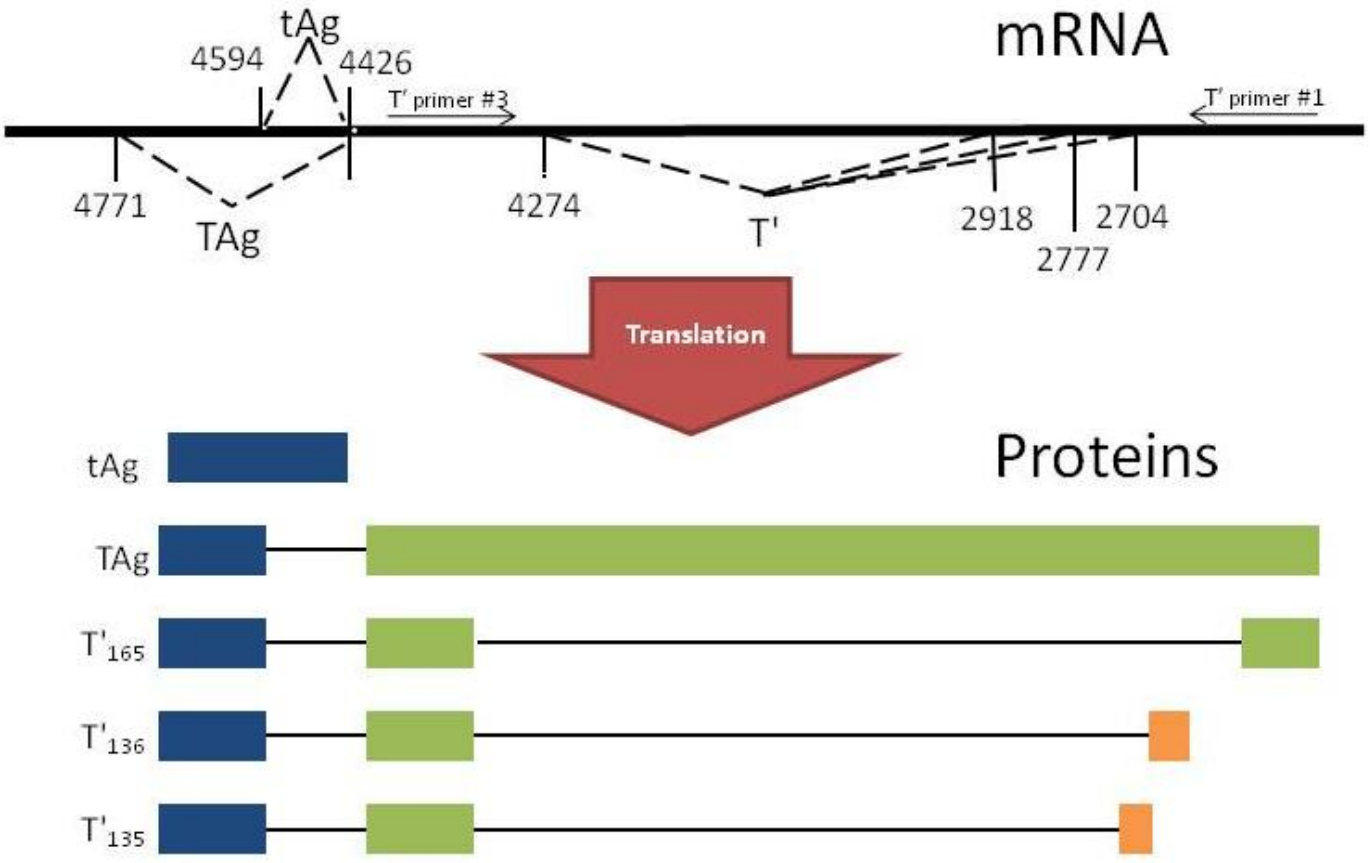

Figure 2. JC Virus Early mRNA and Protein Structures. Shown above are the five early mRNAs produced via an alternative splicing mechanism and the corresponding proteins. After transcription from the early DNA template strand, early mRNA is alternatively spliced to yield TAg, tAg, T´135, T´136, and T´165 transcripts. The key splice sites are labeled on the mRNA strand. The three T´ mRNAs share the same first donor and acceptor sites with large T antigen, and thus have the same initial 132 amino acid protein sequence. T´ mRNAs have a common second donor splice site at nucleotide 4274, and unique second acceptor sites at 2704 (T´165), 2777 (T´136), and 2918 (T´135). The location of T´ primers #1 and #3 utilized in the RT-PCR protocol are indicated on the mRNA strand above.

In their 1995 publication that describes the discovery of the three T´ variants, Trowbridge and Frisque discussed a reverse transcriptase polymerase chain reaction (RT-PCR) experiment in which RNA was extracted from JCV-infected cells and reverse transcribed into cDNAs. After amplification by PCR, bands representing the five cDNAs encoding the JCV early T proteins were

visualized via polyacrylamide gel (PAG) electrophoresis (See Figure 3). Based upon the RT-PCR primers used (Primers #1 and #3, See Figure 2), it was expected that the TAg and tAg cDNAs would be the same size, and therefore form a single band positioned above the three T´ bands. Curiously, a doublet appears in this position on the gel, and no investigations have conclusively demonstrated the presence of TAg or tAg sequences in either of these two bands (See Figure 3). To unravel this mystery, I have attempted to clone and identify the origin of these sequences within the cDNA doublet bands. I have taken a number of approaches to identify these cDNA bands, but I have repeatedly detected only the much smaller T'_{136} cDNA in my experiments.

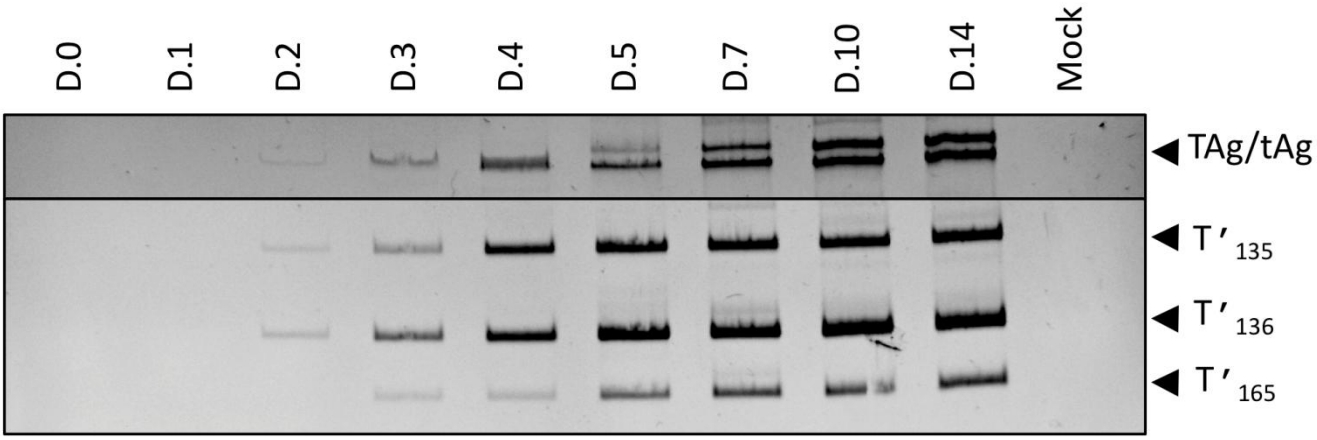

Figure 3. Polyacrylamide Gel of cDNA Bands. Shown above is an image of a polyacrylamide gel with five cDNA bands representing proteins encoded by the early JCV region. In this RT-PCR time course experiment, cDNA was reverse transcribed from mRNA extracted over fourteen days from mammalian cells infected with JC virus. The five bands described in previous publications can be observed at increasing intensities over time, with the greatest amount of cDNA product seen at day fourteen. The two bands investigated during my project are labeled above as TAg/tAg, and are referred to as the "doublet" bands throughout this paper. (Experiment performed by Ramya Kartikeyan, Kartikeyan and Frisque, 2011).

Materials and Methods

RT-PCR

mRNA analysis was performed through a two-step RT-PCR process. Total mRNA from JCV transformed POJ-19 cells or from infected mammalian cells was extracted using Trizol (Invitrogen). cDNA was synthesized from 1 μg of mRNA treated with DNase I (Promega) to remove contaminating DNA. The DNased RNA was then used to synthesize cDNA using the RevertAid First Strand cDNA Synthesis kit (Fermentas). All PCR reactions were performed using the 2X DreamTaq Green PCR Master Mix (Fermentas) in a thermal cycler (MyCycler; Bio-Rad Laboratories). PCR products were resolved on 7% polyacrylamide gels at 100V for 1 hour. Primers used to amplify the JCV early region were T´ primer #1 (JCV 2580-2602; 5´-CCAGCTTTACTTAACAGTTGCAG -3´) and T´ primer #3 (JCV 4368-4345; 5´-GGGATGAAAGACCTGTTTTGCCATG-3´) {Trowbridge and Frisque, 1995}. In some experiments a nested PCR procedure was carried out using extracted cDNA product from the previous PCR reaction as the template strand. Alternative T´ primers #3 (5´-TGTTTGCCAGTGATGATGAAAACACA-3´) and #1 (5´- GGGACCCAGACATGATGAGA3´) were used for this procedure.

DNA Extraction from Polyacrylamide Gels

Polyacrylamide gels were stained with ethidium bromide for 20 minutes. The desired gel bands were visualized by long wave UV illumination and excised. An equal volume of elution buffer (0.5 ammonium acete, 1mM EDTA pH 8.0) was added to the tube, which was then incubated at 37°C overnight on a rotating wheel. Two protocols were used to purify these DNAs. A phenol/chloroform extraction was followed for the first samples (including A1). After overnight incubation, the liquefied gel bands were centrifuged at 12,000 rpm for 10 minutes. The isolated

supernatants were treated with an equal volume of phenol and spun in a microfuge. The aqueous top later was collected and treated with an equal volume of a chloroform solution (24:1 $ChCl_3$:Iso-amyl alcohol). After additional centrifugation, the upper aqueous layer was removed, the DNA solution brought to a concentration of 0.15M NaCl, and a 2X volume of absolute ethanol added. The solution was chilled at -70°C for 15 minutes and then microfuged for an additional 15 minutes in a cold room. Excess ethanol was decanted, and the pellet was dried by lyophilization and suspended in 50µL dH_20. DNA yields were high using this procedure, but difficulties were encountered with ligation and bacterial transformation efficiency, possibly due to chloroform contamination. To avoid the phenol-chloroform extraction step, subsequent DNA purifications were carried out using the Wizard® SV Gel/PCR Clean-Up Kit System (Promega Corporation, Madison, WI).

Ligation

Ligation of the purified DNA was performed using the CloneJET™ PCR Cloning Kit (Fermentas Life Sciences Corporation). Purified PCR product was used at a 3:1 molar ratio with the pJET1.2/blunt cloning vector (50ng/µl). For the first set of samples (A1), the Blunt-End Cloning Protocol was followed. A ligation reaction containing the cloning vector (1µL), PCR product, and T4 DNA ligase (5u/µl), was incubated at room temperature (22°C) for 30 minutes. The remaining PCR product samples underwent the Sticky-End Cloning Protocol, to ensure that any 3´-dA overhangs potentially generated by the polymerase would not interfere with the blunt-end ligation reaction. Briefly, a blunting reaction was performed by treating the PCR product solution with a blunting enzyme for 5 minutes at 70°C. To this reaction mixture was added 1µL each of the cloning vector and T4 DNA ligase, and then incubated for 30 minutes at room temperature. A

control ligation product was created using the Blunt-End protocol and the 976 bp control insert provided in the cloning kit.

Transformation of Escherichia coli

Ten µL of ligation product were added to 100µL of Z competent *E. coli* (DH5α) cells, which were then incubated on ice for 20 minutes. SOC medium (400µL; room temperature; 2% bactotryptone, 0.5% yeast extract, 10mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2 \cdot 6H_20$, 10mM $MgSO_4 \cdot 7H_20$, 20 mM glucose) was added to the bacteria and the solution was shaken at 37°C. After 1 hour, 200 µL of the bacterial mixture was plated on a Luria-Bertani (LB) ampicillin plate (1% bactotryptone, 0.5% yeast extract, 1% NaCl, 1.5% bacto agar, 1% ampicillin) and incubated overnight at 37°C. Individual colonies were selected and grown overnight in LB broth (1% bactotryptone, 0.5% yeast, 1% NaCl) with ampicillin. Extraction of plasmid DNA from the *E. coli* cells was performed using the GeneJET™ Plasmid Miniprep Kit (Fermentas). Fifty ng of extracted product was electrophoresed on a 1.2% agarose gel to screen for evidence of a DNA insertion in the clones. Samples and a control ligation product and 5Kb ladder, were electrophored for 30 minutes at 100V.

Restriction Enzyme Digestion

To further investigate whether the purified plasmid clones contained a DNA insert, restriction endonuclease digestion was performed. For the first set of samples (A1), 1µ BamHI (20U/µL) was added to a reaction mix containing 1µL of purified clone DNA. Restriction enzyme BamHI was chosen because its recognition site (5´-GGATCC-3´) is present in the JCV early coding region but not in the pJET1.2 vector. After 1 hour of digestion at 37°C, cleaved products were electrophoresed on a 1.2% agarose gel for 20 minutes. Undigested samples were also included on

the gel as controls. Another digestion approach was taken with clone samples derived later. To visualize the entire insert band, enzymes XhoI (5´-CTCGAG-3´) and XbaI (5´-TCTAGA-3´), with recognition sites located in the MCS region on either side of the PCR insert, were used (See Figure 4). For each sample, 150ng of DNA were mixed with 0.5µL XbaI (10U/µL) and 0.5µL XhoI (15U/µL) and incubated for 2 hours at 37°C. Cleaved products were then run versus uncleaved products on a 1.2% agarose gel for 15-20 minutes at 100V.

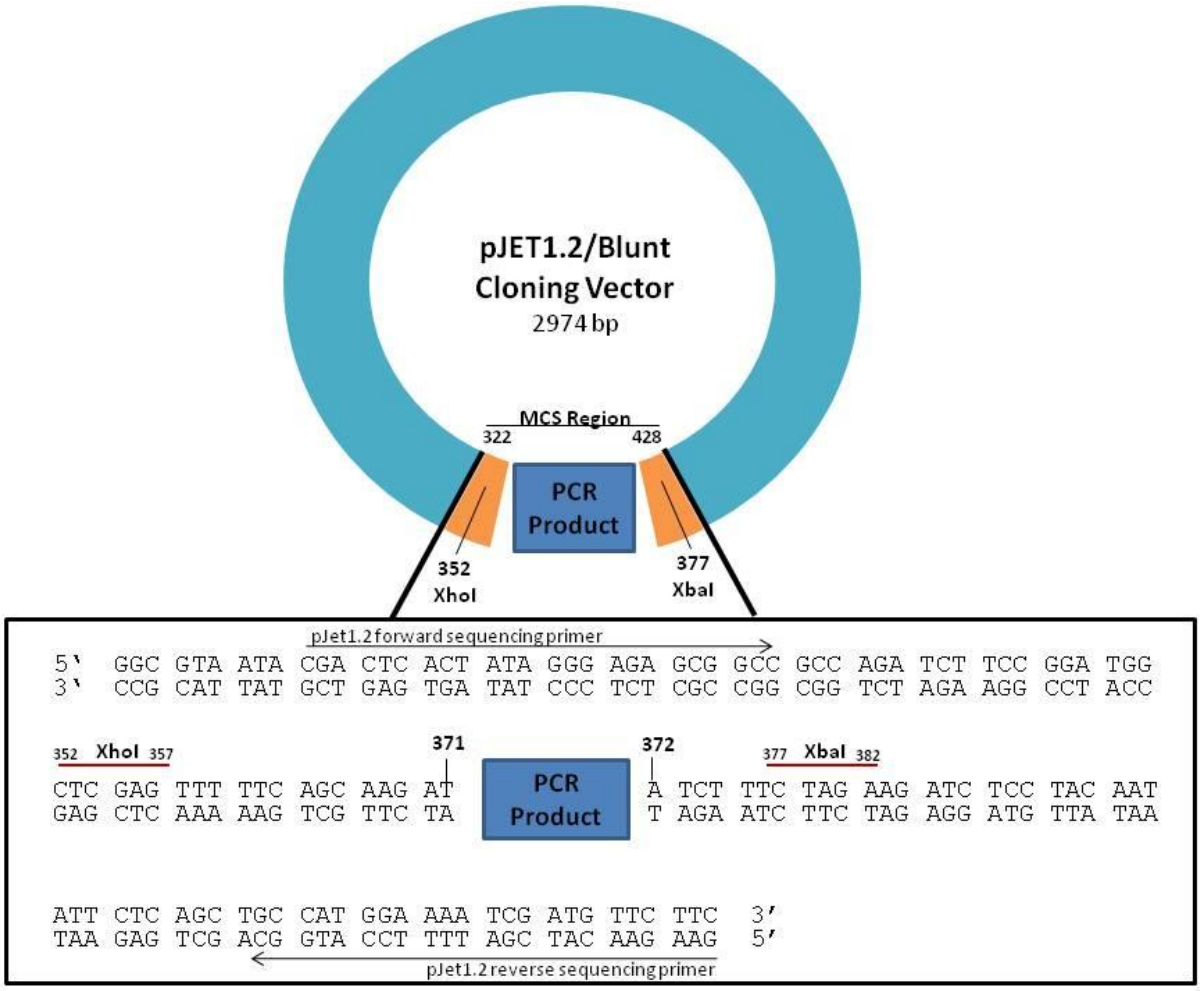

Figure 4. Diagram of pJET1.2/blunt cloning vector. The 2974 bp vector contains a multiple cloning cite (MCS) region. This sequence, spanning nucleotides 322-428 has many restriction enzyme sites, including XhoI and XbaI, which were utilized in this project. Blunt-end PCR products ligate between vector nucleotides 371 and 372. Forward and reverse primer sequences are labeled with arrows. This cloning vector was provided in the CloneJET™Kit (Fermentas).

DNA Sequencing

Select cloned cDNAs were chosen for further screening by DNA sequencing at Penn State's Nucleic Acid Facility. For each sample, 5µL of plasmid DNA (200-300ng/µL), 5µL of pJET1.2 forward sequencing primer (1µM), and 5µL of pJET1.2 reverse sequencing primer (1µM) from the CloneJET™ PCR Cloning Kit (Fermentas) were provided to the facility for sequence analysis.

Results

First Trial of JCV Early cDNA Characterization Yields a T´$_{136}$ and T´$_{165}$ Clone

To investigate the identity of the JCV cDNAs visualized on polyacrylamide gels following RT-PCR amplification of JCV early mRNAs extracted from infected cells, DNAs were removed from the gels and analyzed through a series of procedures. RT-PCR was performed using primers that target the JCV early coding region including TAg and tAg. Polyacrylamide gels separated the amplified DNA by size, with a doublet band clearly forming above three faster migrating bands (See Figure 3). Based upon the primers used, the cDNAs of TAg/tAg, T´$_{165}$, T´$_{136}$, and T´$_{135}$ should be 1788, 432, 291, and 218 nucleotides long, respectively. The top two bands were excised separately from the polyacrylamide gel. From each band, DNA was extracted, purified, and ligated into a blunt cloning vector to allow for transformation in *E. coli* cells. Bacterial colony growth was observed only for cells transfected by the DNA product from the uppermost of the doublet bands. Four of these colonies were picked for overnight culture, from which DNA was purified using a miniprep kit. DNA was quantified by nanodrop and visualized by gel electrophoresis.

Endonuclease digestion with BamHI was performed to investigate the presence of JCV sequences. This enzyme linearized the cDNA sample plasmids that contain a JCV insert. If the intended JCV insert of TAg/tAg (1788 bp) ligated correctly into the 2974 bp vector sequence, a plasmid of approximately 4762 bp would be expected. However, the bands representing cleaved DNA only appeared to be slightly greater than 3000 bp in length, suggesting that a TAg/tAg cDNA insert was not present. Despite this, one sample was sent for sequencing to determine the identity of a short insert that appeared to be present in one plasmid. Sequencing results unexpectedly revealed that two JCV sequences were present. Using T´ primer # 3, the JCV sequence read from 4368, the 5´ end of the primer, to 4274, the unique donor splice site for the 3 T´ transcripts. At this

point, the sequence appeared to be scrambled, but in studying the sequencing chromatograph, we realized that the sequence represented two different clones with sequences starting either at nucleotide 2777 or 2704 (See Figure 5). Nucleotide 2777 is located at 3´ acceptor site for the T'_{136} transcript, and nucleotide 2704 is the location of the 3´ splice site for the T'_{165} transcript. The overlapping sequences indicate two cDNA clones consisting of T'_{136} and T'_{165} sequence were present in my sample.

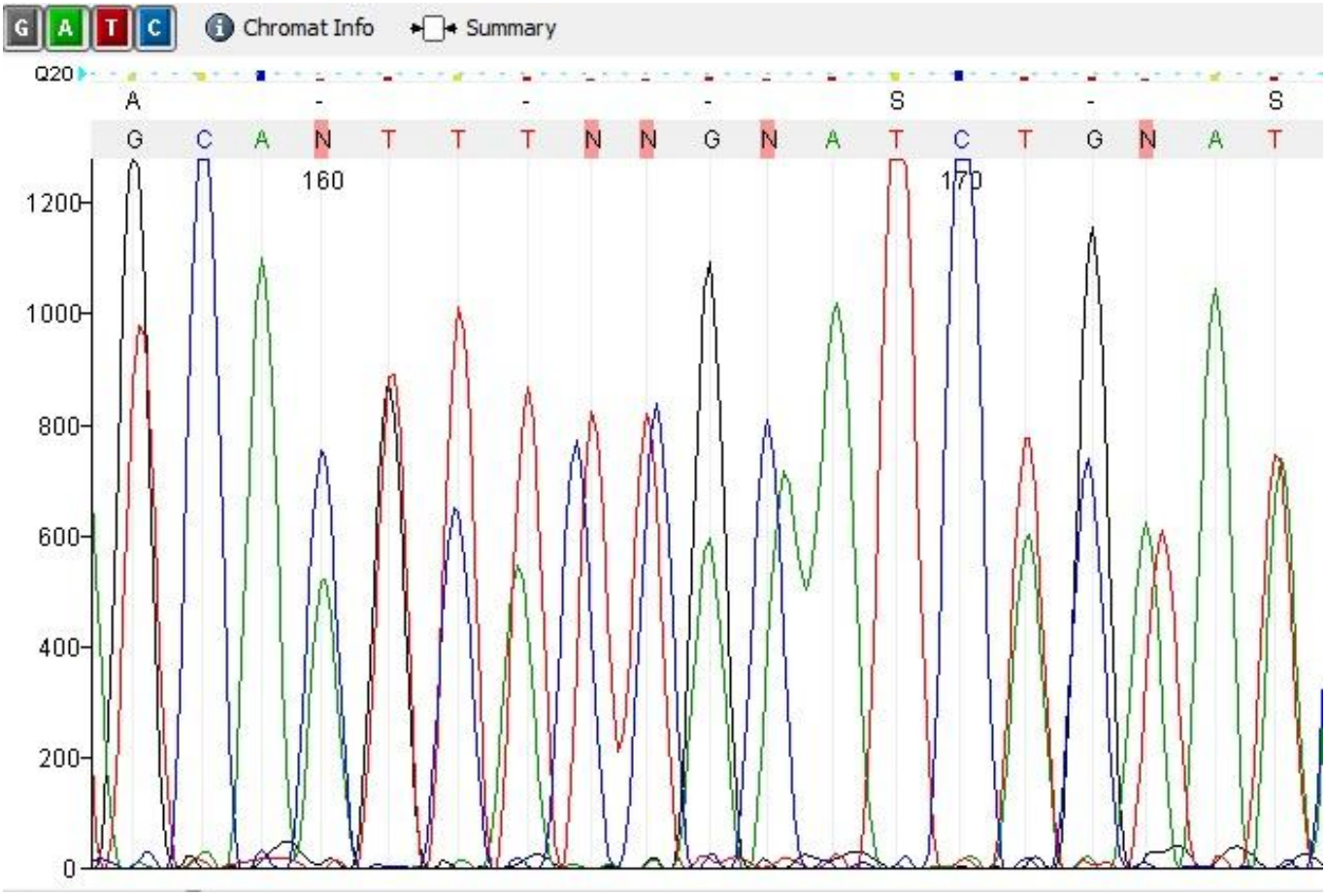

Figure 5. Sequencing Chromatogram. Seen above is the sequencing data output in the form of a fluorescent peak trace chromatogram for sample A1 between nucleotides 157 and 175. Automated DNA sequencers determines the nucleotide bases – adenosine (A, green), guanine (G, grey), cytosine (C, blue), and thymine (T, red) – based upon dye fluorescence biotechnology after capillary electrophoresis. Clear chromatogram results show evenly-spaced, single color peaks and minimal baseline "noise". The chromatogram for A1 has two overlapping peaks throughout this part of the sequence. We determined that the dual peaks signaled the presence of both the T'_{136} and T'_{165} cDNAs.

Second Trial of JCV Early cDNA Characterization Detects Presence of T´$_{136}$ Only

Problems with low DNA yields in the isolation and purification steps in the first trial led to the adjustment of several protocols for the next set of experiments. Utilizing the same T´ primers, a new RT-PCR experiment using RNA extracted from JCV-infected PHFG cells at day 14 post-infection was performed. Two approaches were taken to process the amplification products. First, 50µL of the RT-PCR reaction mix were purified using the Wizard® SV Gel/PCR Clean-Up Kit (Promega) and ligated without undergoing band separation by gel electrophoresis (a shotgun cloning approach). Since T´ sequences were found in the top band in the first trial, the shotgun ligation was an attempt to ligate all of the cDNAs present in the RT-PCR solution, which should include the TAg /tAg coding sequences as well as those of the T´s. However, several unsuccessful ligation attempts led me to try another avenue of experimental methods.

The remaining 200µL of unpurified RT-PCR product were electrophored on a polyacrylamide gel, and the doublet cDNA bands were excised together to avoid DNA loss or damage. Additionally, the middle of three lower cDNA bands, known to represent T´$_{136}$, was excised separately to serve as a control. From each band, DNA was extracted, purified, and ligated into a blunt cloning vector. Ligation product was used to transform *E. coli* cells. Colony growth was observed for cells transfected with either doublet and T´$_{136}$ band cDNA products, and individual colonies were picked for overnight culture. Extraction of DNA from E. coli cells was carried out using a miniprep kit, and nanodrop analysis was performed to calculate purified DNA concentrations. DNAs were visualized by gel electrophoresis and were compared to the control ligation plasmid (3950 bp) to estimate their sizes. Bands running parallel to or above the control DNA were chosen for enzyme digestion.

To screen the cloned ligation products for the presence of the expected insert, a double digest with XbaI and XhoI was performed. Rather than linearizing vector containing JCV DNA by

digestion with a single enzyme (i.e. BamHI), this method removed the intact JCV insert, permitting better estimation of its size. Electrophoresis of the digestion products showed that XbaI and XhoI both cleaved pJET1.2 vector, and that the ligation procedure had been successful in creating plasmids containing an insert. However, these inserts were again much smaller than anticipated, as observed when pairs of uncleaved and cleaved DNA samples were run alongside a control DNA on an agarose gel (See Figure 6). The control product, generated by using the CloneJET™ PCR Cloning Kit, was a 3950 bp plasmid (976 bp insert and 2974 bp vector). After restriction enzyme digestion, the control insert is seen migrating significantly higher on the gel compared to the experimental insert bands, indicating my clones had inserts smaller than 976 bp. It had been expected that the experimental inserts would migrate at a position closer to 1800 bp on the gel.

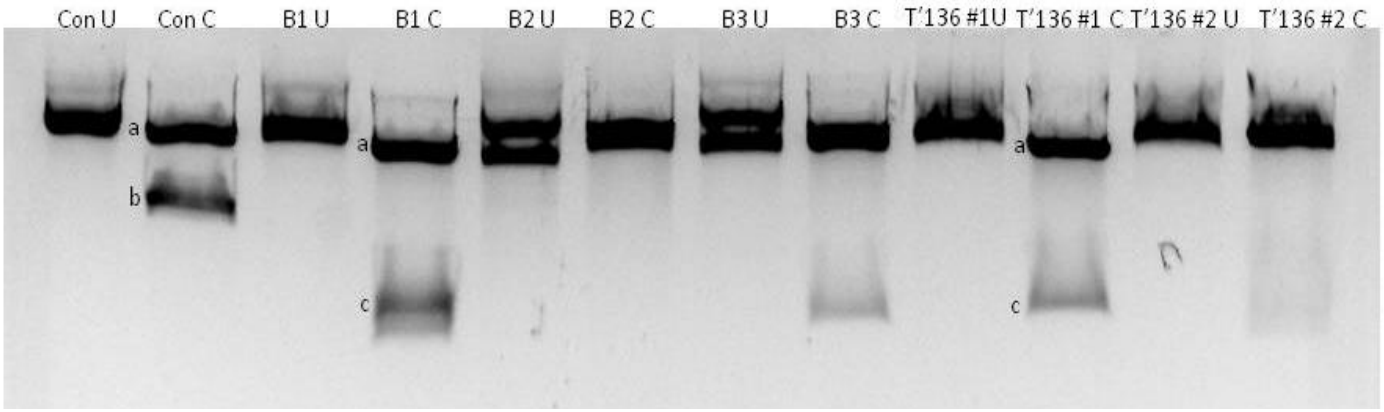

Figure 6. Restriction Digest 1. Restriction digest of purified sticky-end ligation product from bacterial clones containing PCR product. Restriction enzymes XbaI & XhoI used to cleave pJET1.2 vector sequence on either side of the DNA insert. Insert bands visualized on 1.2% agarose gel after 150ng purified ligation product was treated with enzyme mixture for two hours at optimal reaction temperature. PCR template sources from top and bottom (B samples) and T´136 polyacrylamide gel bands. A DNA sample was either electrophored on the gel as linear DNA cleaved with XbaI + XhoI (C) or as undigested supercoiled DNA (U). The ligation control sample (Con) is included for analysis of experimental samples and to verify ligation efficiency. Cleaved recombinant plasmids present two bands representing the linearized vector (a) at 2974 bp and an insert. The control insert (b) is 976 bp, and PCR product inserts (c) are approxminately 310 bp. Experimental sample inserts were predicted to run higher compared to the control insert since they were expected to be longer (~1800 bp) in nucleotide length.

A clone from the doublet band (B1) and T'_{136} band (T'_{136} #1), were selected for further study because they contained inserts and were originally generated from different sized cDNA bands. Sequence analysis of both DNAs revealed the two DNAs were identical and represented the T'_{136} cDNA inserted into the cloning vector. Using T´ primers #3 and #1, the JCV sequence was found to span nucleotides 4368 to 2580. Sequences between nucleotides 4274 (shared T´ donor site position) and 2777 (T'_{136} acceptor splice site) were absent from the clones as expected for a T'_{136} cDNA. This T'_{136} sequence is 291 nucleotides in length, which is slightly shorter than the total insert size of 310 bp due to the addition of nucleotides from the vector after the XbaI and XhoI restriction sites. The presence of splice variant T'_{136} in the B1 clone was unexpected and failed to support our hypothesis that one of the doublet bands represents the TAg and tAg cDNAs. The presence of T'_{136} sequence in the control clone did confirm, as expected, that the middle of three lower cDNA bands contains T'_{136} cDNA.

To examine further the possibility that some clones prepared from the doublet cDNA bands contained TAg and tAg sequences, additional bacterial colonies obtained from this experiment were processed following the same methods. Samples were initially screened for the presence of an insert after purification of plasmid DNAs from bacterial clones. Undigested samples were electrophored on agarose gel versus the control plasmid and a 5Kb DNA ladder (See Figure 7).

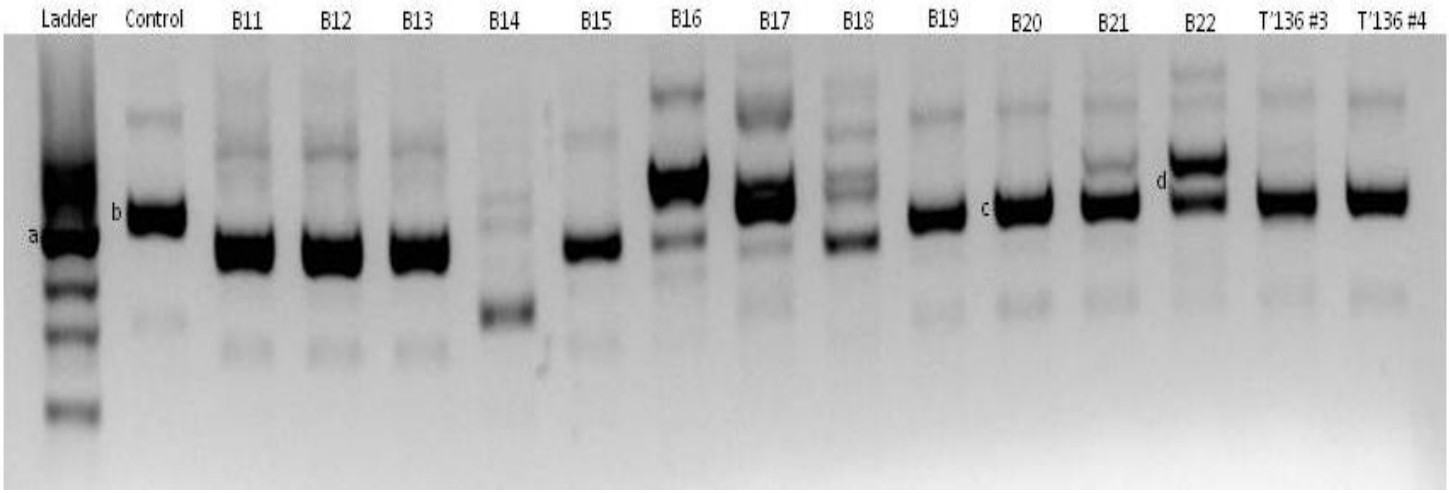

Figure 7. Clone Screening. Purified clones were screened for the presence of RT-PCR products inerted into plasmid by electrophoresis on a 1.2% agarose gel. DNA samples purified by a miniprep kit were run versus 5kb linear DNA ladder and the control sample. Ladder bands represent linear DNA sizes, with the band labeled as (a) marking 3000bp linear DNA. The supercoiled ligation control sample (b), which is known to be 3950bp, was used as a benchmark for size comparisons with the minipreped JCV samples. Supercoiled recombinant plasmids, such as sample B20 (c), that migrated at the same position or higher on the gel relative to the control were chosen for further analysis by restriction enzyme digestion. Double bands may indicate nicked circle DNA (d), where one of the two DNA strands is broken and the nicked circular DNA migrates slower on a gel compared to supercoiled DNA.

Ladder bands mark linear DNA size, with the band denoted by "a" indicating 3000 bp. However, the control plasmid at 3950 bp (b) was used as a reference for comparing sizes of the experimental clones due to their shared supercoiled DNA structure. Structural features of DNA affect how a DNA migrates during electrophoresis, making comparisons of linear to supercoiled DNA of the same size on agarose gels misleading. Samples migrating at the same position or above the control DNA (B16, B17, B20, B22) were suspected of having an insert, and were chosen for further analysis. As indicated in Figure 7, observation of closely migrating bands in the same sample likely indicates nicked circles, another conformational variation of plasmid DNA characterized by one nicked strand.

Restriction enzyme digestion was performed on four samples of interest, B16, B17, B20, and B22. Uncleaved and cleaved clone samples electrophoresed in an agarose gel can be seen in Figure 8. The sample inserts ran at a similar position/size as those of the previous restriction digest #1 (Figure 6), and clone B20 was chosen for sequencing.

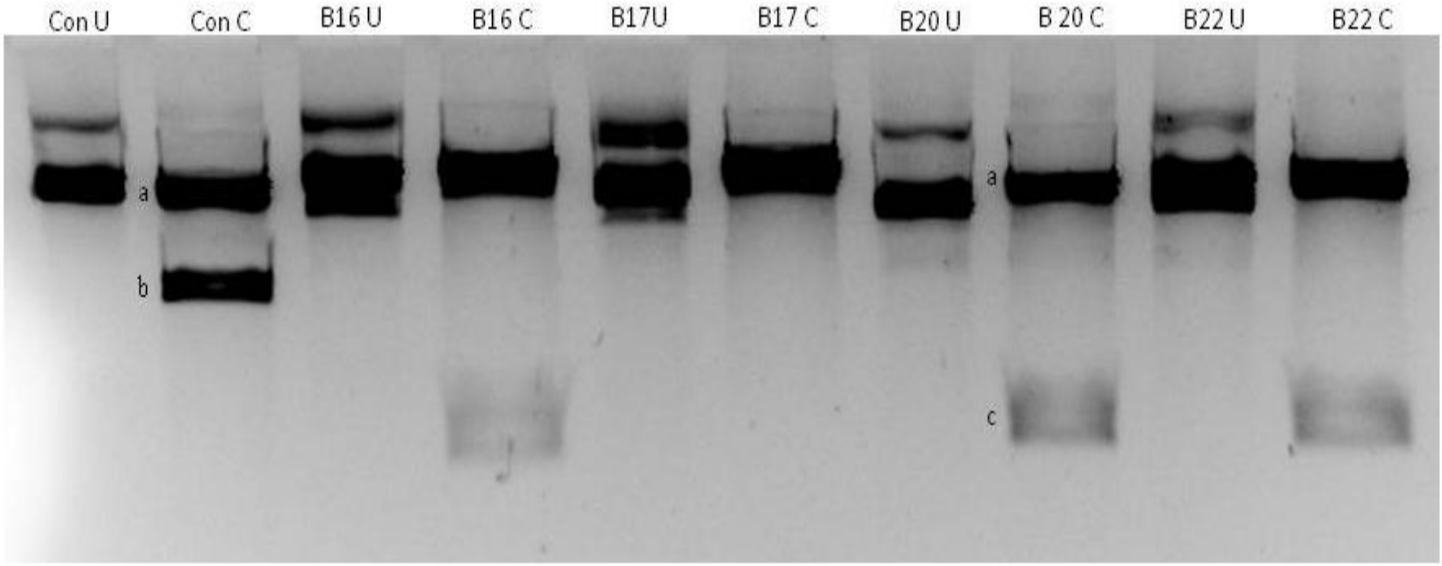

Figure 8. Restriction Digest 2. For a detailed description see Figure 6. Cleaved recombinant plasmids yield two bands representing the linearized vector (a) at 2974 bp and an insert. The control insert (b) is 976 bp, and RT-PCR product inserts (c) are approximately 310 bp.

This same experimental procedure was used to screen another set of 20 clones containing ligated RT-PCR product (screening data not shown). After restriction enzyme analysis, sample B32 was sent for sequencing with B20. The results again showed the presence of a T′$_{136}$ insert, mirroring the results of the previous experiments. Figure 8 presents the sequencing analysis data for sample B32. The *N* at position 364 in the initial pJET1.2 vector sequence represents a nucleotide that could not be determined by the sequencing equipment, and represents a sequencing artifact that occurs commonly near the primer binding site. The correct nucleotide at that position in the vector sequence is an A. Another slight inconsistency in the early vector sequence is the absence of two A's; where the sequence reads 5′-GCAG-3′, it should be 5′-GCAAAG-3′. The deletion of these two nucleotides at this location in the vector sequence was observed in every clone.

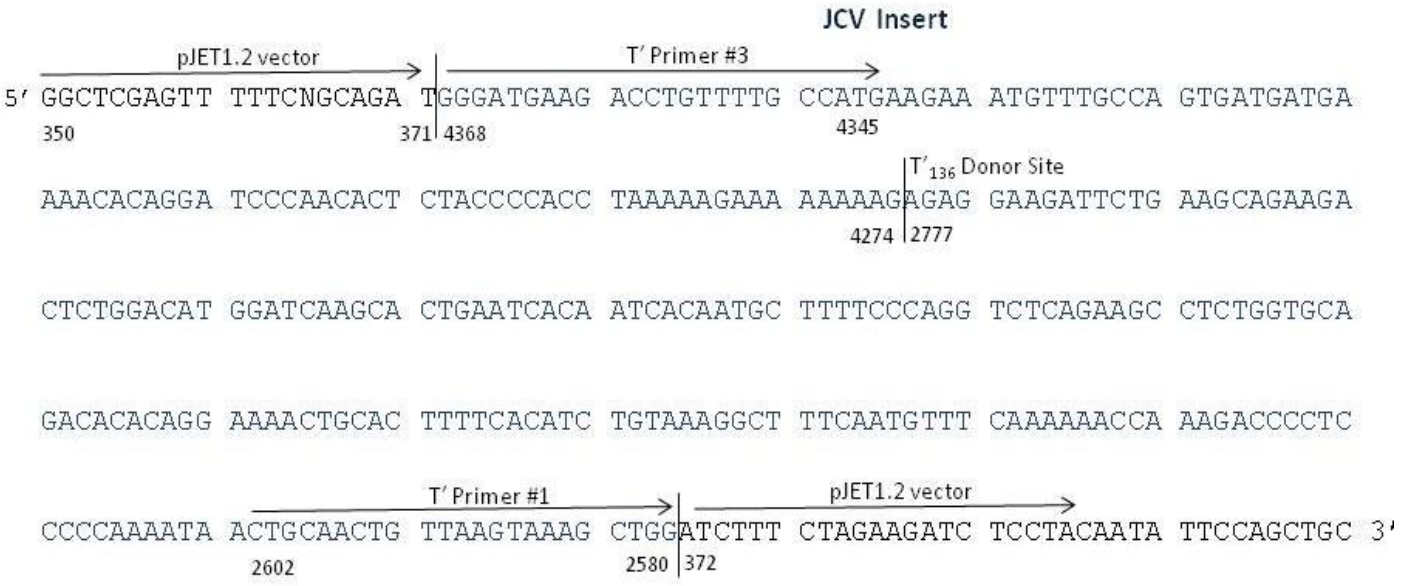

Figure 9. Sequencing Diagram. A schematic showing sequencing results for sample B32. A single DNA band is presented in the 5' to 3' direction. Locations of the cloning vector and JCV insert sequences are labeled with numbered start and end nucleotides. The T′ Primer sequences marked with arrows indicate the boundaries for cDNA amplification during RT-PCR. The JCV insert sequence reads from the beginning of the T′ Primer #3 until the shared T′ donor splice site at 4274. Intron sequence between 4273 and 2778 are removed and the JCV sequence resumes at 2777 at the start of the third exon of T′$_{136}$. The rest of the JCV sequence reads until the end of the T′ Primer #1 at 2580.

Third Trial Using Nested PCR to Isolate Doublet Band cDNAs for Amplification

The repeated detection of the T'_{136} gene in bands predicted to represent the TAg/tAg cDNA was difficult to explain. It seems likely that the doublet bands were contaminated by the smaller, faster migrating T'_{136} cDNA. To enhance amplification of predominant TAg/tAg cDNAs, a second PCR step was conducted using the purified cDNA generated by the first RT-PCR amplification. The same RT-PCR amplification protocol was followed using alternate versions of T´ primers #1 and #3, whose sequences lie directly inward of the original primers. While these primers still target the same early JCV region, it was necessary to use different primers to avoid potential sequence degeneracy at the original primer sites.

A shotgun ligation approach was taken since the amplified cDNA should be solely from the top two bands of interest. To investigate which purification and ligation methods worked most efficiently, several combinations of techniques were used to process the PCR product. Raw PCR product and PCR product purified by Wizard™ PCR Clean-Up Kit were each ligated in separate Sticky- and Blunt-End reactions. Thereby, four sets of bacterial clones were made by different purification/ligation methods; Unpurified/Sticky-End, Purified/Sticky-End, Unpurified/Blunt-End, and Purified/Blunt-End. The Unpurified/Sticky-End approach produced the greatest number of bacterial colonies, but restriction enzyme analysis revealed low ligation efficiency, with no significant inserts observed.

Analysis of the other three clone types showed evidence of insert ligation, as seen in Figure 10. While insert bands can be faintly seen for the Unpurified/Blunt-End (C3, C4) and Purified/Blunt-End (C5) samples, the bands from the Purified/Sticky-End clones (C1, C2) were of greatest interest. In prior digests, experimental insert bands ran consistently at the same position as b1. Band b2 appeared notably higher on the gel (though still not as high as predicted), potentially indicating that clone C2 contained an insert different than the T'_{136}

sequence previously observed. Results for C1 and C2 provided two novel sequencing patterns, though both have been identified as T'_{136} due to the unique acceptor splice site at 2777. Sample C1 has a shortened sequence (4338-4274, 2777-2635) compared to prior T'_{136} inserts. The sequence begins at 4338, the alternate T′ primer #3 start, and reads to the first donor site at 4274. Continuing at the acceptor site for T'_{136} (2777), the sequence reads until a premature end at nucleotide 2635. Neither #1 T′ Primer (original or alternate) sequence is reached, raising a question as to how the cDNA became truncated. Even stranger is the sequencing result for sample C2. Starting four nucleotides short of the T′ Primer #3 at 4334, the sequence actually continues past the original T′ primer #1 (2580) and ends at nucleotide 2498. Since the T'_{136} donor splice site is present, it can be presumed that this sample is too, a cDNA of $T'_{136.}$ It is unclear, however, as to how a cDNA that is longer than its template strand could be produced using PCR primers targeting an abbreviated sequence.

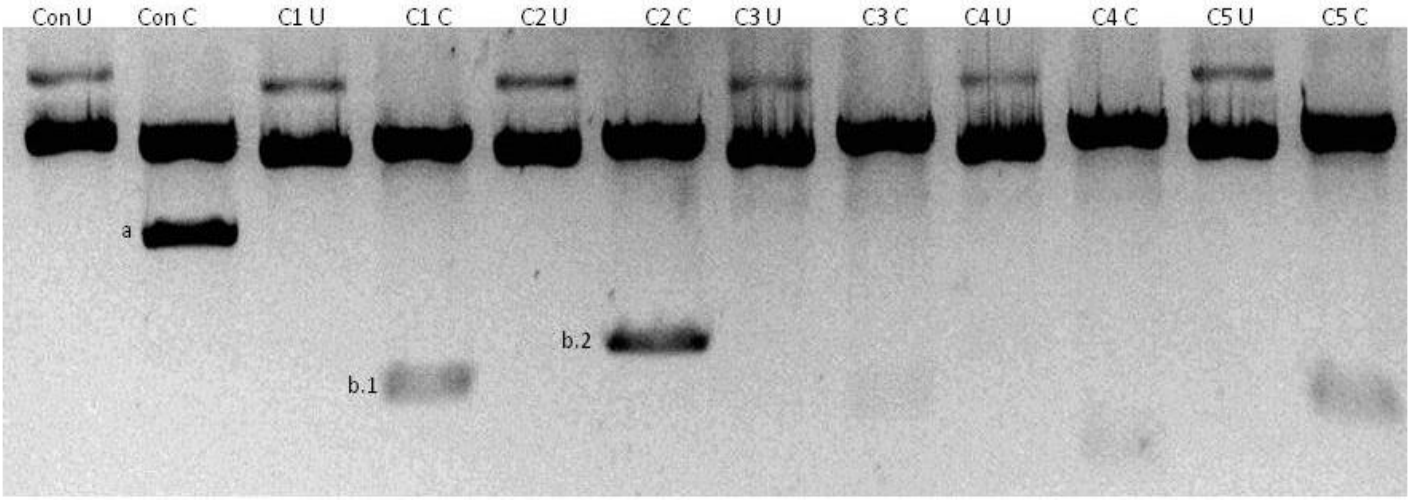

Figure 10. Restriction Digest 3. Clones of purified and ligated cDNA from the nested PCR were digested with restriction endonucleases XbaI and XhoI. Uncleaved (U) and cleaved (C) samples were electrophoressed in adjacent lanes on the gel. Insert lengths (b) were estimated by comparison with the known control insert (a). Inserts for samples C1 (b.1) and C2 (b.2) migrated at different rates on the gel, indicating DNA inserts of different lengths. Clones from three different purification/ligation methods are seen above, Purified/sticky-End (C1, C2), Unpurified/Blunt-End (C3, C4), and Purified/Blunt-End (C5). Samples C1 and C2 yielded bands of different sizes, and were sequenced to determine their identity.

Through the experimental process, I have consistently detected the T'_{136} splice pattern within the cDNAs isolated from doublet bands. In total, seven clones were analyzed for the presence of JCV sequence, six from doublet band cDNA and one from the T'_{136} band cDNA. While different methods were utilized for sample production throughout this experiment, all clones showed at least a partial sequence of T'_{136} (Table 1). Four clones (B1, B20, B32, and T''_{136} #1) all shared an identical, intact T'_{136} sequence.

Table 1. Summary of Results

Clone Names	Source of cDNA	Ligation Technique[a]	JCV Sequence	Clone Identity
A1	Excised Top Band	Blunt End	4368-4274, [b] 2777-2580/2704-2580	T′136 & T′165[c]
T′136 #1	Excised T'_{136} Control Band	Sticky End	4368-4274, 2777-2580	T′136
B1	Excised Top & Bottom Bands	Sticky End	4368-4274, 2777-2580	T′136
B20	Excised Top & Bottom Bands	Sticky End	4368-4274, 2777-2580	T′136
B32	Excised Top & Bottom Bands	Sticky End	4368-4274, 2777-2580	T′136
C1	Nested PCR - T&B Bands	Sticky End	4338-4274, 2777-2635	T′136
C3	Nested PCR - T&B Bands	Sticky End	4334-4274, 2777-2498	T′136

Footnotes

[a] JCV inserts were cloned using Fermentas CloneJET PCR Cloning Kit. Cloning protocols were provided for Blunt End and Sticky End ligation methods. Blunt-end PCR products can be directly ligated into the pJET1.2/blunt cloning vector. PCR products with 3'-dA overhangs ("Sticky End" products) must be blunted with a DNA blunting enzyme prior to ligation.

[b] The sequencing chromatograph for Clone A1 showed an overlap of two JCV sequences after the first donor splice site at 4274. One represented T'_{136} sequences, reading from 2777-2580. The overlapping nucleotide sequence represented T'_{165} sequences of 2704-2580.

[c] The overlapping sequence indicated the presence of both T'_{136} and T'_{165} clones.

Discussion

Researchers in the mid-1990s performed RT-PCR analysis of JCV early mRNA, expecting to produce four cDNAs. Instead, five bands were observed on a polyacrylamide gel, an unexplained finding that has remained a mystery to date. For my Honors thesis project, I chose to identify the mystery band. mRNA extracted from JCV-infected mammalian cells was reverse transcribed to form a cDNA template for PCR. An alternative splicing mechanism removes introns from immature early region mRNA to create five unique mRNAs, TAg, tAg, $T'_{135,}$ T'_{136} and T'_{165}. Despite being the smallest protein, tAg actually has the longest mature mRNA sequence. An early stop codon at 4494 prevents further translation past the splice acceptor site at 4426. The shared TAg/T´ intron (4771-4426) removes this stop codon, allowing translation to continue to nucleotide 2603 for the TAg transcript. The T´ mRNAs each have a unique second intron that is removed to form three different sized mature mRNAs. After PCR amplification, cDNAs derived from these transcripts produce five separate bands on polyacrylamide gels. However, only four bands were expected that would represent four cDNA sizes. The primers used in RT-PCR analysis, T´ primers #3 and #1, limit the boundaries for cDNA amplification to JCV nucleotides 4368 though 2580, respectively. TAg and tAg have identical sequences in this region because the introns by which they differ lie outside the region amplified by the RT-PCR primers. Thus, the cDNAs of TAg and tAg produced during PCR amplification should be of identical size, and should yield a single gel band of approximately 1800 bp, however, two bands (referred to here as the doublet bands) are observed near this position on the gel.

My investigation into this question began by excising each band in the doublet separately in order to isolate and identify each cDNA. Difficulties encountered in the ligation and transformation processes resulted in only a few bacterial colonies, and these were only generated from the uppermost band cDNA. DNA sequencing revealed the presence of two different

sequences, T'_{136} and T'_{165}, in one clone. This was an unexpected result for two reasons, the first being the detection of two species of cDNA from a single clone. Transformation of ligated cDNA product forms bacterial colonies derived from single recombinant plasmids, and therefore are expected to contain a single DNA sequence. A possible explanation for the detection of overlapping T'_{136} and T'_{165} sequences is that two colonies were accidentally picked together and grown in one overnight culture, leading to the observation of two, not one, DNA sequences. However, additional confusion resulted from the identification of T′ splice variant cDNAs within the band predicted to contain TAg/tAg cDNA. The T'_{136} and T'_{165} cDNA sequences are 291 and 432 nucleotides in length, respectively, and are found in separate, faster migrating bands (See Figure 3). Detecting these significantly shorter T′ cDNAs in the uppermost doublet band thought to contain cDNA approximately 1800 bp in length challenged the basic principles of the RT-PCR and PAG techniques used in this experiment. These observations led to adjustments in experimental procedures to try to detect the TAg/tAg cDNA.

Throughout the next several trials, different experimental approaches were attempted to investigate which would maximize ligation efficiency and DNA yield, and ultimately lead to the characterization of TAg/tAg cDNAs. A shotgun ligation of unpurified PCR product was conducted to bypass the many purification steps that could potentially cause DNA loss and create recombinant plasmids containing all cDNAs amplified by RT-PCR. Taking this approach we obtained a number of transformed bacterial colonies, but after multiple rounds of screening these colonies by restriction enzyme digestion and agarose gel electrophoresis, it did not appear that any of the recombinant DNA clones contained the target 1800 bp insert. Returning to the PAG approach, the doublet bands were excised once again to isolate the cDNA from the lower doublet band which had not yet yielded bacterial clones in the initial experiment. The two bands were excised together this time to try to reduce DNA loss or damage that may occur during the

extraction procedure. A Sticky-End ligation protocol was followed to ensure that any potential overhangs resulting from the PCR amplification step would not reduce the ligation efficiency. Over forty clones were screened from this trial, and the four clones sequenced all contained the T'_{136} sequence. The repeated detection of T'_{136} confirmed that this sequence is, in fact, present in the doublet bands, but raised questions regarding the unexpected size and the absence of a TAg/tAg insert. The 291 bp sequence could not solely comprise the doublet bands because cDNA of that size would have a faster migration profile on the gel. It was suspected that a small amount of the lower molecular weight cDNA was becoming "trapped" amongst the larger TAg/tAg cDNA during electrophoresis, preventing some of the T'_{136} cDNA from properly migrating through the polyacrylamide gel.

Though finding T'_{136} in the doublet bands was unanticipated, I still suspected that TAg/tAg cDNA was the dominant species within these bands, and, for some reason, was not being detected using our experimental methods. In my final experimental trial, cDNA purified from previously excised doublet bands was used as the template for a second RT-PCR amplification. This way, only cDNA sequences extracted from physically isolated doublet bands would be "reamplified", with the expectation that this approach would increase my chances to detect the predominant sequence, TAg/tAg, over T'_{136}. To optimize results, four combinations of various purification and ligation techniques were used to produce clones, though no significant increase in bacterial transformation efficiency was observed with any specific approach. Enzyme digestion analysis showed that one clone, C2, contained an insert of a greater length than any observed in previous experiments (See Figure 10, b.2). This insert was smaller than the 976 bp control sequence, and therefore still not likely to be the predicted 1788 bp TAg/tAg cDNA, but the possible presence of a different insert sequence was intriguing. Another clone, C1, had an insert that appeared to migrate at a rate similar to samples from

previous trials that proved to be T´$_{136}$ (See Figure 10, b.1). Through DNA sequencing analysis, it was concluded that these two clones were both T´$_{136}$ due to the presence of the unique T´$_{136}$ donor-acceptor splice sites. Clone C1's slightly shorter T´$_{136}$ sequence was expected due to the alternate primers utilized in the second PCR. Strangely, Clone C2 contained a T´$_{136}$ sequence that continued past both alternate and original T´ primer #1 sequences and ended at nucleotide 2498. This explains the appearance of a larger insert during gel electrophoresis, however, it is unclear how this sequence was produced. The alternate T´ primers used for the second PCR bind directly inward of the original T´ primers, resulting in the amplification of a slightly shorter cDNA sequence. It is possible that these alternate primers did not fully bind to the complementary sequence on the template strand, allowing for amplification of sequence past the primer boundaries. However, the template sequence for the second PCR was a cDNA amplified in a prior PCR using the original T´ primers #3 and #1, so that the template strand itself should only represent the DNA fragment from nucleotides 4368 to 2580. The observation of a cDNA sequence produced by a nested-PCR procedure that continues past both sets of primer sequences is a puzzling result that may be due to inproper binding of both the alternate and original T´ primer #1 to JCV sequences downstream from the intended target template sequence.

During the course of my experiments, I repeatedly detected the presence of the smaller T´$_{136}$ cDNA in large molecular weight bands that I expected would contain TAg/tAg cDNA. It is possible that a portion of the shorter T´$_{136}$ cDNA molecules became trapped in the slower migrating TAg/tAg cDNA, contaminating the top bands. If such an event did occur, the TAg/tAg sequence should still predominate over any other cDNA sequences present, since the doublet band must contain cDNAs significantly larger than the T´ sequences based on band-migration rates. My consistent detection of T´$_{136}$ indicates that my experimental methods preferentially selected the smaller T´ sequence that, though less abundant in the doublet bands,

was more efficiently cloned.

Only relatively small DNA fragments are ligated into plasmid vectors; the pJET 1.2/blunt cloning vector used in this experiment is recommended for cloning inserts from 6 bp to 10,000 bp. While these restraints easily include the intended 1788 bp TAg/tAg insert, DNA size can affect ligation efficiency. In a ligation mixture containing both TAg/tAg and T'_{136} sequences, I predict that the vector would ligate the 291 bp T'_{136} cDNA with greater efficiency than the 1788 bp TAg/tAg cDNA. Longer DNA sequences also are more susceptible to damage by UV light during gel excision, potentially resulting in low DNA quality and low bacterial transformation yield.

Apart from the ligation technique, another possible explanation for my inability to clone the TAg/tAg cDNA is that this JCV clone sequence contains elements toxic to bacteria. Certain DNA sequences, especially in the presence of strong *E. coli* promoters, transcribe the production of toxic proteins not tolerated by bacteria. The longer cDNA sequence from the doublet bands might contain one of these antibacterial elements, causing toxicity in the *E. coli* cells and the loss of bacteria containing a vector with the TAg/tAg sequence. Further troubleshooting by optimizing temperature, growth serum, or vector sequence conditions may allow for cloning of our target sequence in the future.

After characterization of cDNAs from the PAG doublet, it remains unclear if TAg/tAg is found in either of the bands. In addition, I also do not know the identity of the cDNA in the second of the doublet bands. It is possible that alternative splicing of the JCV early pre-mRNA yields a sixth JCV transcript that has gone undetected to date. Alternatively, a host mRNA may be recognized by our RT-PCR primers and the second doublet band contains amplified cellular DNA sequences. Once clones are obtained that contain the correct sized insert, sequence analysis should unravel these mysteries.

References

1. Allander, T., K. Andreasson, S. Gupta, A. Bjerkner, G. Bogdanovic, M. A. Persson, T. Dalianis, T. Ramqvist, and B. Andersson. 2007. Identification of a third human polyomavirus. *J Virol.* **81**:4130-4136.
2. Bollag, B., C. Prins, E.L. Snyder, R.J. Frisque, 2000. Purified JC virus T and T´ proteins differentially interact with the Retinoblastoma family of tumor suppressor proteins. *Virology* **274**:165-178.
3. Bollag, B., L.H. Kilpatrick, S.K. Tyagarajan, M.J. Tevethia, and R.J. Frisque. 2006. JC virus $T´_{135}$, $T´_{136}$, and $T´_{165}$ proteins interact with cellular p107 and p130 *in vivo* and influence viral transformation potential. *J. Neurovirol*. **12**:428-442.
4. Cui, X., J.C. Wang, A. Deckhit, B.C. Joseph, P. Eberwein, C.L. Cubitt, C.F. Ryschkewitsch, H.T. Agnostini, and G.L. Stoner. 2004. Chinese strains (Type 7) of JC Virus are Afro-Asiatic in origin but are phylogenetically distinct from the Mongolian and Indian strains (Type 2D) and the Korean and Japanese strains (Type 2A). *J. Mol. Evol*. **58**:568-583.
5. Cho, U.S., S. Morrone, A.A. Sablina, J.D. Arroyo, W.C. Hahn, and W Xu. 2007. Structural basis of PP2A inhibiton by small t Antigen. *PLoS Biol*. **5**:1810-1819.
6. Elsner, C and K. Dorries. 1992. Evidence of human polyomavirus BK and JC infection in normal brain tissue. *Virology* **191**: 72-80.
7. Feng, H., M. Shua, Y. Chang, and P.S. Moore. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319**:1096-1100.
8. Frisque, R.J. 2001. Structure and function of JC virus T´ proteins. *J. Neurovirol*. **7**:293-297.
9. Frisque, R.J., B. Bollag, S.K. Tyagarajan, and L.H. Kilpatrick. 2003. T´ proteins influence JC virus biology. *J. Neurovirol.* **9**:15-20.
10. Frisque, R.J.,G.L. Bream, and M.T. Cannella. 1984. Human polyomavirus JC virus genome. *J.*

Virol. **51:** 458-469.

11. Gardner, S. D., A. M. Field, D. V. Coleman, and B. Hulme. 1971. New human papovavirus (B.K.) isolated from urine after renal transplantation. *Lancet* **i**:1253-7.
12. Gaynor, A. M., M. D. Nissen, D. M. Whiley, I. M. Mackay, S. B. Lambert, G. Wu, D. C. Brennan, G. A. Storch, T. P. Sloots, and D. Wang. 2007. Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog* **3**:e64.
13. Giacinti, C., and A. Giordano. 2006. RB and cell cycle progression. *Oncogene* **25**:5520-5227.
14. Hutchinson, M. 2007. Natalizumab: a new treatment for relapsing remitting multiple sclerosis. *Ther. Clin. Risk Manag*. **3**: 259-268.
15. Jeong, B.H., K. Lee, E. Choi, K. Kim, and Y. Kim. 2004. Genotyping of the JC Virus in urine samples of healthy korean individuals. *J. Med. Virol*. **72**: 281-289.
16. Johnson, E.M. 2010. Structural evaluation of new human polyomaviruses provides clues to pathobiology. *Cell Press* **15**: 215-223.
17. Kartikeyan R, and R.J. Frisque. Y-box cellular factor 1 and viral agnoprotein influence regulation of alternative splicing of JC virus early mRNA. *Manuscript in preparation (to be submitted in April 2011).*
18. Mori, M., H. Kurata, M. Tajima, H. Shimada. 1991. JC virus detection by in situ hybridization in brain tissue from elderly patients. *Ann Neurol*. **29**:428-432.
19. Padgett, B.L., D.L. Walker, G.M. ZuRhein, R.J. Eckroade, and B.H. Dessel. 1971. Cultivation of papova-like virus from human brain with progressive multifocal leukoencephalopathy. *Lancet* **i:** 1257-1260.
20. Prins, C, and R.J. Frisque. 2001. JC virus T´ proteins encoded by alternatively spliced early mRNAs enhance T antigen-mediated viral DNA replication in human cells. *J. Neurovirol.* **7**:250-264.

21. Reviriego-Mendoza, M.M., and R.J Frisque. 2011. Interaction and co-localization of JC virus large T antigen and the F-box protein β-transducin-repeat containing protein. *Virology* **410**:119-128.

22. Schowalter, R. M., D. V. Pastrana, K. A. Pumphrey, A. L. Moyer, and C. B. Buck. 2010. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe.* **7**:509-515.

23. Sunyeav, SR, A. Lugovsky, K. Simon, L. Gorelik. 2009. Adaptive mutation in the JC virus protein capsid are associated with progressive multifocal leukoencephalopathy. *PLoS Genet*. **5**: e1000368.

24. Trowbridge, P.W., and R.J. Frisque. 1995. Identification of three new JC virus proteins generated by alternative splicing of early viral mRNA. *J. Neurovirol*. **80**:2083-2091.

25. van der Meijden, E., R. W. Janssens, C. Lauber, J. N. Bouwes Bavinck, A. E. Gorbalenya, and M. C. Feltkamp. 2010. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathog.* **6**:e1001024.

26. Walker, D.L., B.L. Padgett, G.M. ZuRhein, A.E. Albert, and R.F. Marsh. 1973. Human papovavirus (JC): induction of brain tumors in hamsters. *Science* **181**:674-676.

Appendix A.

Glossary

capsid: the protein coat of a virus that protects the viral genome

coding region: the region of DNA sequence that is transcribed and processed into a mature mRNA, and. therefore, codes for protein

domain: a part of a protein that has a distinct tertiary structure; often attributed to distinct biological function

exon: the sequences in a gene that remain in a mature messenger-RNA (mRNA) and, therefore, code for the protein

genotype: genetic variants in a population

intron: the sequences in a gene that are not include in a messenger-RNA (mRNA) because they are removed during the RNA splicing process

in vitro: taking place in a cell-free solution, e.g., in a test tube

in vivo: taing place within a living cell or organism

lytic infection: a viral infection that results in the rupture of the cell membrane of the host cell, releasing the new virus particle produces in that cell

molecular chaperone: proteins that function to maintain other proteins in the correct folded state, or conformation

plasmid: a piece of circular, extra-chromosomal bacterial DNA with the ability to independently replicate in a cell

polymerase chain reaction: use of a thermostable DNA polymerase enzyme to replicate and amplify a DNA template in vitro and using a repetitive serious of temperature variations

splice variants: proteins that are related to one another because they result from translation of transcripts that have been produced, from the same precursor-mRNA by alternative splicing

transformation: process by which a cell acquires characteristics consistent with those of a cancer cell

transcription factor: a protein that regulates transcription, the process by which messenger-RNA is synthesized from a gene in the first step necessary for the expression of the protein encoded by that gene

Appendix B.

Academic Vita

ERIN S BARNES

927 Oakwood Ave. Spring Grove, PA 17362
erinsbarnes@psu.edu

EDUCATION:

The Pennsylvania State University

- Eberly College of Science
- B.S. in Biology– Vertebrate Physiology Option, Spring 2011
- Honors in Biochemistry and Molecular Biology

THESIS TITLE: Detection and Characterization of Viral cDNAs Amplified from Cells Expressing the JC Virus Early Coding Sequences.

THESIS SUPERVISOR: Dr. Richard J. Frisque, Professor of Molecular Virology

WORK EXPERIENCE:

Undergraduate Research at Penn State University June 2008- May 2011
Biochemistry & Molecular Biology Lab Position - P.I. Dr. Richard Frisque

- Research project on the Oncogenic and Pathogenic Potential of Human Polyomavirus (more specifically the JC virus)

University of Colorado Cancer Fellowship June 2010 – Aug 2010
Medical Oncology Summer Research Fellow – Dr. Virginia Borges MD

- Research project investigating the immunological implications of novel clinical breast cancer treatments
- Abstract publication on the *Function of Myeloid Derived Suppressor Cells Across the Biologic Subtypes of Breast Cancer*.

GRANTS RECIEVED:

- 2008 Undergraduate Summer Discovery Grant

SCHOLARSHIPS AND AWARDS:

- Schreyer Honors College Academic Excellence Award
- Safari Club International Foundation Conservation Scholarship Award
- Spring Grove Area Scholarship Foundation Award
- Dean's List (Fall 2007, 2008, 2009, 2010, Spring 2008, 2010)