2017

Analysis of Variable Regions in Feline Leukemia Virus Envelope Sequences

Brianna Sohl

Follow this and additional works at: http://commons.emich.edu/honors

Recommended Citation
http://commons.emich.edu/honors/536
Analysis of Variable Regions in Feline Leukemia Virus Envelope Sequences

Abstract
Feline Leukemia Virus (FeLV) is a retrovirus that affects domesticated cats and other species in the Felis genus. There are three major subtypes, each characterized by unique disease progression and proviral sequences. FeLV-A is weakly pathogenic, -B has been associated with tumors, and -C causes severe anemia. Proviral sequences were amplified using PCR with FeLV-specific and adapter-specific primers. The method of primer walking has yielded consensus sequences for several fragments of 1,000-3,000 base pairs of proviral DNA. Sequences for ten different clones of the provirus envelope region from a FeLV positive blood sample were compared to known sequences and revealed novel insertions.

Degree Type
Open Access Senior Honors Thesis

Department
Chemistry

First Advisor
Dr. Heather Holmes

Second Advisor
Dr. Hedeele Evans

Keywords
provirus, retrovirus, felis catus, cat

Subject Categories
Chemistry

This open access senior honors thesis is available at DigitalCommons@EMU: http://commons.emich.edu/honors/536
ANALYSIS OF VARIABLE REGIONS IN FELINE LEUKEMIA VIRUS ENVELOPE SEQUENCES

By

Brianna Sohl

A Senior Thesis Submitted to the
Eastern Michigan University
Honors College

In Partial Fulfillment of the Requirements for Graduation

With Honors in Chemistry

Approved at Ypsilanti, Michigan, on this date April 19, 2017

Supervising Instructor (Dr. Heather Holmes)

Department Head (Dr. Deborah Heyl-Clegg)

Honors Director (Dr. Rebecca Sipes)
Abstract

Feline Leukemia Virus (FeLV) is a retrovirus that affects domesticated cats and other species in the *Felis* genus. There are three major subtypes, each characterized by unique disease progression and proviral sequences. FeLV-A is weakly pathogenic, -B has been associated with tumors, and -C causes severe anemia. Proviral sequences were amplified using PCR with FeLV-specific and adapter-specific primers. The method of primer walking has yielded consensus sequences for several fragments of 1,000-3,000 base pairs of proviral DNA. Sequences for ten different clones of the provirus envelope region from a FeLV positive blood sample were compared to known sequences and revealed novel insertions.
Introduction

Feline Leukemia Virus (FeLV) is a gammaretrovirus that affects domestic cats (Felis catus) and closely related species which causes hematological diseases such as lymphoma, leukemia, and severe anemia \[1-3\]. It is widely believed that FeLV originated from a trans-species infection of a rodent inflicted with a gammaretrovirus that, after consumption of the mouse by the ancestral cat, resulted in FeLV \[4\]. Other species that have recently been found to be infected with FeLV include the Florida panther and the endangered Iberian lynx \[4\]. FeLV has also been of significant platform for human research as well; the T-cell lymphomas described in FeLV-positive cats inspired US biomedical researchers to investigate human T-cell leukemia, and eventually these studies culminated in the discovery of human immunodeficiency virus type 1 \[4\].

There are primarily three different major subgroups of FeLV that are differentiated by the receptor the virus uses to enter the cell as well as its host range; the most common subgroup is FeL V-A \[1, 4, 5\]. FeLV-A is responsible for horizontal transmission between animals and gives rise to other FeLV subgroups \[4, 5\]. FeLV-A utilizes the thiamine transporter THTR1 to integrate into host cells \[4\]. It was previously thought that FeLV-A subgroups could be generalized as weakly pathogenic, but recent research into sequence variation of FeLV-A has discovered that subtle mutations have brought about a strain, FeLV-945, of FeLV-A that produced greater pathological consequences for infected cats \[3\].

The two other major subgroups of FeLV include FeLV-B and FeLV-C. FeLV-B is associated with tumor formation and lymphoma, and arises from recombination events between FeLV-A and endogenous FeLV that is already present in the host genome \[5\].
FeLV-B utilizes the phosphate symporter Pit-1 and the closely related Pit-2 as receptors for viral integration \[^4\]. It is widely believed that FeLV-A is necessary for transmission of FeLV-B and FeLV-C, but there has recently been an example of host transmission of FeLV-B in the absence of FeLV-A \[^6\]. FeLV-C is associated with red blood cell aplasia and severe anemia that is fatal within weeks, and binds to the heme transporter FLVCR1 for viral integration into the host cell \[^4\]. The exact genesis of the FeLV-C subgroup is unknown, but is thought to be arisen through recombination or possibly through mutational changes in FeLV-A \[^4\]. There are two other subgroups of FeLV, FeLV-T and FeLV-D, that have yet to be well defined; FeLV-T has been found to be similar to FeLV-AIDS and is thought to arise via mutation, and FeLV-D arose through recombination with FeLV-A and endogenous FeLV, similar to that of FeLV-B \[^5\].

FeLV is a simple retrovirus and therefore only encodes for genes that are required for its structure and replication \[^1\]. The three viral genes that make up the FeLV genome are the \textit{gag}, \textit{pol}, and \textit{env} genes (Figure 1) \[^2\]. The \textit{gag} gene encodes for structural proteins including the capsid of the virus, the \textit{pol} gene encodes for replicating enzymes, and the \textit{env} gene encodes for the surface glycoprotein (SU) and the transmembrane protein (TM) which make up the viral envelope \[^2\]. The differences in subtype therefore arise through mutations within the \textit{env} region, and these mutations are often caused \textit{in vivo} during FeLV replication due to the low-fidelity of reverse transcriptase \[^1\]. The envelope is also responsible for the entry of the retrovirus into host cells as the SU protein, attached to the TM, interacts with host cell surface receptors \[^5\].
**Figure 1: Viral Envelope Structure of Surface Glycoprotein.** The structure of the integrated FeLV provirus, including the viral genes (gag, pol, env) and the long terminal repeat (LTR). An enhanced section of the env gene is brought forward to illustrate the location of Variable Region A (VRA) and Variable Region B (VRB) that construct the RBD. Differences within these regions affect cellular tropism and pathogenicity as the virus integrates into the host cell.

Within the SU protein are two highly variable domains that are integral for interaction and entry to host cells: Variable Region A (VRA) determines receptor binding and host range, and Variable Region B (VRB) is required for efficient infection \[^{12}\]. Adjacent to these two variable regions is the Proline Rich Region (PRR) that mediates conformation changes required for viral entry \[^{2}\]. These three domains together are referred to as the Receptor Binding Domain (RBD) \[^{2}\]. Variations then in the RBD of the SU protein are then of considerable interest as these mutations are thought to influence the rate of viral infection and therefore the resulting pathology of the disease \[^{2}\].

Since FeLV is a simple gammaretrovirus, and therefore only encodes for genes required for structure and replication, there are no genes within the genome of FeLV that are directly malignant \[^{4}\]. Instead, the association of FeLV with tumors and malignancies
arises from the integration of proviral DNA into the host genome and activating oncogenes \(^4\). This oncogene activation is not necessarily targeted by the proviral integration, but certain genomic loci of host cells have been determined to be common insertion sites in which FeLV integrates and causes the activation of oncogenes at a rate greater than random chance \(^4\). FeLV also may undergo a process referred to as retroviral transduction in which endogenous oncogene sequences are incorporated into the viral genome and are then able to infect other host cells through FeLV replication and integration \(^4\).

The goal of this project was to identify mutational changes in the \(env\) gene from the blood of a FeLV-positive cat. Proviral sequences from a FeLV-positive kitten were amplified using PCR with FeLV-specific and adapter-specific primers. The method of primer walking has yielded consensus sequences for several fragments of 1,000–3,000 base pairs of proviral DNA. Through this, novel substitutions and insertions were found in the proviral clones that were located within the PRR and VRB of the RBD that may consequently have affected pathogenicity of the FeLV virus.
Materials and Methods

Extraction of Genomic DNA from Blood Samples

A Nucleospin Tissue Genomic DNA Extraction Kit (Clontech) was used to extract blood and tissue samples.

Preparation of Blood Lysates

Blood samples were lysed by pipetting 25 µL proteinase K and 200 µL whole or EDTA anti-coagulated blood into 1.5 mL microcentrifuge tubes. A volume of 200 µL Buffer B3 was added to the samples, which were then vortexed vigorously (10–20 s). Following incubation at room temperature for 5 min, the samples were vortexed gently, and then incubated at 70 °C for 10-15 min.

Binding, Washing, and Eluting DNA from Blood Lysates

Blood and tissue lysates were vortexed briefly after incubation, then µL ethanol (96–100 %) were added. Samples were then vortexed vigorously, and transferred to a NucleoSpin® spin column and collection tube. After centrifuging for 1 min at 11,000 x g, the flow-through was discarded, and 500 µL Buffer BW were added. The tubes were again centrifuged for 1 min at 11,000 x g, and the flow-through was discarded. An aliquot of 600 µL Buffer B5 was added, the columns were again centrifuged for 1 min at 11,000 x g, and the flow-through was discarded. The columns were then centrifuged for 1 min at 11,000 x g to remove residual ethanol. The dry spin columns were then placed in clean collection tubes. Elution Buffer BE was pre-incubated at 70 °C, then a volume of 50 µL was added to each sample. After incubation at room temperature for 3 minutes, samples were centrifuged for 1 min at 11,000 x g to collect DNA eluent. The spin columns were placed in clean collection tubes, and a second aliquot of 50 µL Elution Buffer BE was
added to each sample. After 3 minutes at room temperature, the samples were again centrifuged for 1 min at 11,000 x g. All samples were then stored as recommended at -20 °C until used.

**Preparation of DNA Libraries**

**Restriction Enzyme Digestion**

DNA templates were digested with restriction enzymes DraI, EcoRV, PvuII, and StuI according to the Genome Walker manual (Clontech). For each digestion, 25 µL of DNA extract were mixed with 8 µL of restriction enzyme (10 units/µL), 10 µL of buffer (10x), and 57 µL of deionized water. The tubes were inverted several times to mix, and incubated at 37 °C for 2 hours. After vortexing gently for 10-15 s, the reactions were incubated overnight at 37 °C.

**Post-Digestion Purification**

A NucleoSpin® Gel and PCR Clean-up Kit (Clontech) was used to purify the digested DNA samples. Spin columns were placed into collection tubes (2 mL) and 500-700 µL of sample digest were loaded. The column was centrifuged for 30 s at 11,000 x g, and the flow-through was discarded. A volume of 700 µL Buffer NT3 was added, and the column was centrifuged for 30 s at 11,000 x g. The flow-through was discarded and the column centrifuged again for 1 min at 11,000 x g to remove Buffer NT3 completely, and were then incubated for 5 min at 70 °C. The columns were then placed into new 1.5 mL microcentrifuge tubes, and 30 µL Buffer NE were added. After incubating at room temperature for 1 min, tubes were centrifuged for 1 min at 11,000 x g, and the eluent was collected. Samples were stored at -20 °C until used.
Ligation of Adapter Sequences to Purified DNA Digests

Adapters were ligated to each sample of purified DNA digest using reagents from and according to the instructions in the Genome Walker kit (Clontech). The reaction mixture was composed of 4.8 µL purified DNA digest, 1.9 µL Genome Walker Adaptor (25 µM), 0.8 µL of 10x Ligation Buffer, and 0.5 µL T4 DNA Ligase (6 units/µL). Samples were incubated at 16 °C overnight in a thermal cycler, then incubated for 5 min at 70 °C to stop the reaction. To each tube, 32 µL of TE buffer (10/1 pH 7.5) were added. The tubes were vortexed gently for 10-15 s, and then stored at -20 °C until used.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>[Mg²⁺] 1 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1R</td>
<td>GGTCCAATTGAGCACCAGCCATGGAAA</td>
<td>65.9</td>
</tr>
<tr>
<td>IB3R</td>
<td>CGGCTAGATAGTGCGCCCCTGTATGT</td>
<td>65.8</td>
</tr>
</tbody>
</table>

PCR Reactions

Master mixes for primary and secondary PCR were prepared by combining reagents as shown in Table 1, for the appropriate number of total reactions. After mixing, 24 µL were added to each PCR reaction tube, along with 1 µL of template, or PCR-grade water for negative controls. Reaction conditions for primary and secondary PCR are shown in Table 2.

Table 1. Reagents for each primary or secondary PCR reaction.

<table>
<thead>
<tr>
<th>Volume (µL)</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.5</td>
<td>PCR-Grade Water</td>
</tr>
<tr>
<td>2.5</td>
<td>10x Advantage 2 PCR Buffer</td>
</tr>
<tr>
<td>0.5</td>
<td>dNTP (10 mM each)</td>
</tr>
<tr>
<td>0.5</td>
<td>Each primer (10 µM)</td>
</tr>
<tr>
<td>0.5</td>
<td>Advantage 2 Polymerase Mix (50x)</td>
</tr>
</tbody>
</table>
Table 2. Reaction conditions for primary and secondary PCR.

<table>
<thead>
<tr>
<th></th>
<th>Primary PCR</th>
<th>Secondary PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td>94°C 25 sec</td>
<td>5 Cycles</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td>72°C 3 min</td>
<td>7 Cycles</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td>94°C 25 sec</td>
<td>20 Cycles</td>
</tr>
<tr>
<td><strong>Step 4</strong></td>
<td>67°C 3 min</td>
<td>5 Cycles</td>
</tr>
<tr>
<td><strong>Step 5</strong></td>
<td>67°C 7 min</td>
<td>1 Cycle</td>
</tr>
</tbody>
</table>

The gene-specific primers used for PCR were designed using NCBI Primer Blast. For primary and secondary PCR, the primer sequences were 5’-GGTCCAATTGAGCACCGCCATGGAAA-3’ and 5’-CGGCTAGATAGTGCGCCCTGTATGT-3’, respectively. Both forward primers were supplied with the Genome Walker kit (Clontech).

Purification of PCR Amplicons

A NucleoSpin® Gel and PCR Clean-up kit (Clontech) was used for purification of PCR products and the extraction of bands from agarose gels.

DNA Extraction from Agarose Gels

A scalpel was used to excise the DNA fragment from an agarose gel, and all excess agarose was removed. The gel slice was weighed and transferred to a clean 1.5 mL tube. For each 100 mg of agarose gel, 200 µL of Buffer NTI was added. The samples were incubated for 5–10 min at 50 °C. The sample was vortexed briefly every 2–3 min until the gel slice was completely dissolved.

Post-PCR Clean-Up

A NucleoSpin® Gel and PCR Clean-up Column was placed into a collection tube.
(2 mL) and up to 700 µL of sample was loaded and centrifuged for 30s at 11,000 x g. The flow-through was discarded and the column was placed back into the collection tube. The remaining samples were loaded and the centrifugation step repeated. A volume of 700 µL Buffer NT3 were added to the column, which was then centrifuged for 30 s at 11,000 x g. The flow-through was discarded and the column placed back into the collection tube. The column was centrifuged for 1 min at 11,000 x g to remove Buffer NT3 completely. The column was then placed into a new 1.5 mL microcentrifuge tube, and 30 µL Buffer NE were added. The column was incubated at room temperature for 1 min, and the tube was centrifuged for 1 min at 11,000 x g to collect the eluent.

**Cloning and Transformation**

Cloning was performed using a TOPO-TA Cloning Kit for Sequencing. Briefly, 2 µL of PCR product was mixed with 1 µL of Salt Solution, 2 µL of water, and 1 µL of pCR-4-TOPO vector. Reactions were incubated at room temperature for 20-30 minutes, then placed on ice for transformation into One Shot DH5α-T1 chemically competent cells. Each vial of cells was thawed on ice for 5 minutes before adding 2 µL of a cloning reaction. Vials were incubated on ice for 20-30 minutes, heat shocked at 42°C for 30 seconds, and placed back on ice for the addition of 250 µL of S.O.C. medium. Tubes were then shaken for 1 hour at 200 rpm and a temperature of 37°C. Colonies were grown on LB plates with either ampicillin or kanamycin, then cultured overnight in LB medium for extraction.

**Plasmid Extraction**

A PureLink® Quick Plasmid Miniprep Kit was used to isolate plasmid DNA. Harvesting was accomplished by centrifuging 5 mL of overnight LB-culture. All medium
was removed. A volume of 250 µL Resuspension Buffer (R3) with RNase A was added to the cell pellet and the pellet was resuspended until it is homogeneous. A volume of 250 µL Lysis Buffer (L7) was added and the solution was mixed gently by inverting the capped tube until the mixture was homogeneous. The tube was incubated at room temperature for 5 minutes, then 350 µL Precipitation Buffer (N4) was added. The solution was mixed immediately by inverting the tube, or for large pellets, vigorously shaking the tube, until the mixture was homogeneous. The lysate was centrifuged at 12,000 × g for 10 minutes.

The supernatant was loaded onto a spin column in a 2-mL wash tube and centrifuged at 12,000 × g for 1 minute. The flow-through was discarded. Another wash was done by adding 500 µL Wash Buffer (W10) with ethanol to the column, allowing the column to incubate at room temperature for 1 minute, then centrifuging at 12,000 × g for 1 minute. The flow-through was discarded, and 700 µL Wash Buffer (W9) with ethanol was added to the column, which was then centrifuged at 12,000 × g for 1 minute. The flow-through was discarded, and the column was centrifuged again 12,000 × g for 1 minute. The spin column was placed in a clean 1.5-mL recovery tube. A volume of 75 µL of preheated (70 °C) TE Buffer (TE) was added to the center of the column, which was incubated for 1 minute at room temperature, then centrifuged at 12,000 × g for 2 minutes to elute the purified plasmid DNA. The plasmid DNA was stored in aliquots at −20°C until used.

**Sequencing**

Samples were submitted for standard Sanger sequencing at the University of Michigan Core Sequencing Lab.
Results and Discussion

Two amplicons containing proviral envelope sequences were isolated from an agarose gel following secondary PCR (Figure 2). The lane in Fig. 2 shows two DNA bands blocked in red and blue boxes of approximately 1000 and 3000 base pairs, respectively. These gel bands were then extracted, cloned and transformed. Five colonies were isolated and sequenced. After consensus sequences were determined using sequencing chromatograms, these clones were then compared to known sequences found through relevant literature and the NCBI website (Figure 3). The alignment in Fig. 3 was constructed using ClustalX software. Refer to Appendix A for all clonal consensus sequences. An important note are the large insertions which are typically characteristic of FeLV-B envelope sequences as recombination events occur between FeLV-A and endogenous FeLV.

It is with interest then, as illustrated by arrows in Figure 3, Clones B1, B2, and B3 had a novel insertion. However, this insertion was not found in any of the other known sequences of FeLV, including FeLV-B. The position of these insertions was found to be located in the PRR of the RBD. This then may have affected the way these particular clones would enter its host cell via the virus’ conformational changes. The resulting efficiency (or inefficiency) of viral
Figure 3: DNA Sequence Alignment of Proviral FeLV Clones and Reference Sequences. Envelope sequences of reference FeLV sequences were used to align with consensus sequences of clonal DNA generated throughout the experiment. FeLV-A: M18248, AYF052723, EU359306, EU359307, EU359308, EU359309, M12500. FeLV-B: K01208, K01209, EU783973, J034481. FeLV-C: M14331. FeLV-T: M87886. FeLV-FAIDS: M18247. Murine Leukemia Virus: Z11128, U94692. Black arrows indicate locations of novel insertions within proviral clones. Asterisk stars indicate consensus base pair across all sequences.
integration and entry based on these insertions could then change resulting pathology of the virus.

Another insertion was detected in Clone B2 within the VRB of the RBD (Figure 3). As illustrated, this insertion was not seen in any of the other clones or known sequences. VRB determines efficient infection of the virus and therefore this insertion may affect this efficiency of resulting infection. However, the insertion also causes a frame-shift mutation, which will likely cause a nonfunctional clone of the FeLV, due to the dramatic change in downstream amino acid sequence.

Due to these novel insertion sites, clones of interest were then compared to predicted protein structures of known envelope sequences of FeLV-A, -B, and -C subgroups (Figure 4). The distinct structure of each envelope structure follows the pattern of each subgroup utilizing different receptors and the usage of receptor differentiation to determine subtype [2]. Though structure may vary slightly, the general structure between strains within the FeLV-A subgroup would likely be extremely similar as FeLV-A has a >97% amino acid similarity within the envelope gene despite geographical location or reverse transcriptase infidelity [2]. Predicted protein structures were constructed using RaptorX software for each clone; RaptorX uses amino acid sequences associated with known protein secondary structure to predict the secondary structure of the submitted input [9]. Clones A1 through A5 were not included in this report due to their small length in amino acid sequence past the envelope open reading frame (ORF) for translation.
Figure 4: Predicted Envelope Protein Structures of Known FeLV Subtypes. (A) The structure of Rickard-A envelope sequence, a FeLV-A strain. (B) The structure of Gardner-Arnstein envelope sequence, a FeLV-B strain. (C) The structure of Sarma envelope sequence, a FeLV-C strain. These three predicted protein structures are included for comparison to notable predicted protein structure of Clones.
Clone B2’s unique insertion was observed in a 3D secondary protein structure (Figure 5). It was compared to FeLV-A, the Rickard strain, due to the close homology of this clonal sequence to the A subgroup. As shown, the insertion caused a loss of secondary structure beta sheet. Due to the insertion’s critical location within the RBD, it is likely that this loss of structure is indicative of a loss of function mutation as well within clone B2.

**Figure 5: Novel Insertion Causes Loss of Secondary Structure.** Above are 3D images of the predicted envelope protein domain I structure of the Rickard A FeLV strain (left) and Clone B2 (right). Red boxes indicate the area in which the insertion was located. As seen in Clone B2, this insertion caused loss of beta sheet structure which may affect binding and efficient infection to host cells.

Clone B3 was selected as a representative to view the predicted protein structure due to its larger consensus amino acid sequence than Clone B1; Clone B2 was not selected due to its frameshift mutation (Figure 6). The arrow within the figure draws attention to the location of the four amino acid insertion site. Though this is in a predicted unstructured area of the protein, it should be noted that this insertion still lies within the PRR and therefore could affect conformational changes of the viral entry as well as antibody/antigen interactions.
After examining individual clonal envelope structure, the evolutionary relationship between all clonal DNA sequences was compared to known sequences by constructing a phylogenetic tree (Figure 7). The outgroups for this phylogenetic tree were the murine leukemia virus and the friend murine leukemia virus which are distantly related to FeLV; the murine leukemia virus is believed to be the ancestral virus from which FeLV developed [4]. All FeLV-B subtypes were grouped together on the bottom branch (Fig. 7) after a divergence event from the FeLV-A subgroup. The FeLV-C strain (M14331) was shown to branch off on its own with slightly less divergence from FeLV-A subtype than the divergence between FeLV-B and FeLV-A. One FeLV-T strain (M87886) and one FeLV-FAIDS strain (M18247) are closely connected to one another which supports the evidence that suggests the close relationship both genetically and pathologically between FeLV-T and FAIDS [5, 8]. The other reference sequences are

**Figure 6: Clone B3 Predicted Envelope SU Protein Structure.** Above is the full predicted protein structure of the envelope of Clone B3. The blue arrow indicates the general location of where the novel insertion occurred. Though there is no noticeable secondary structure, this insertion is within the PRR and can affect conformational changes during viral fusion to host cell. Other changes in PRR sequence have been shown to affect antigen/antibody binding which suggests this insertion to Clone B3 could influence immunological responses from the host against infection.
strains within the FeLV-A subgroup. As shown in Fig. 7, these A subgroup strains have minor divergence events, but remain clustered together. Clones A1-5 and Clones B1-5 however show close relation, but divergence, from all FeLV-A subgroups. Particularly, Clones B1-3, which all include the novel insertion in the PRR, are shown to be most closely evolutionary related to each other out of all 10 total clones. This divergence from the A subgroup may be indicative of a stronger departure from the FeLV-A sequence which typically has >97% amino acid sequence identity. The consequences of this difference could be comparable to that of FeLV-945 which had a greater pathogenic effect on its host cell due to a possible increase in its binding affinity or integration kinetics.

*Figure 7: Phylogenetic Tree of Known FeLV Sequences in Relation to Proviral Clones.* This tree was completed with maximal parsimony using Seaview after alignment with ClustalX. FeLV envelope DNA sequences were randomized five times and bootstrapped 1000 times. Clones were closest in evolutionary relation to the FeLV-A reference sequences. The bottom branch contains FeLV-B subgroup. FeLV-C is scattered throughout as the sequence consistency of the -C subgroup is low. The outgroups used were envelope sequences of the murine leukemia virus and Friend leukemia virus.
Conclusions and Future Directions

Novel insertion sites have been detected in Clones B1-3 that are located within the receptor binding domain of the FeLV envelope region. Due to the location of these insertions, pathology of the resulting viral progression may be affected; this may suggest why Cat 28—the source of the samples used in this project—died at such a young age from a subgroup most closely resembling FeLV-A which is typically weakly pathogenic.

Future directions of this project include screening other blood and tissue samples from the original population at a feline leukemia cat rescue to obtain live virus which will be required for further viral assays. Though Cat 28 died as a kitten, the mother was also FeLV-positive and may have horizontally transmitted the same or very similar strains of FeLV to other cats. If the viral insertions are not found from sample screening of similar cats, an alternative route to obtaining live virus would be to genetically modify the samples collected using CRISPR/Cas9 technology in order to recreate the insertions found in this experiment [10]. Once live virus has been collected, a viral interference assay could then be performed in order to accurately determine the subtype of the FeLV strain [11]. Other future experiments should include a cell binding assay that will analyze viral integration of the clones into host cells to examine binding affinity and kinetics [12]. This will be of interest due to the insertions being located within the receptor binding domain of the FeLV. A neutralizing antibody assay should also be conducted to analyze the possible effects of the insertion on antibody-antigen interactions and therefore immunological response of the infected host [6].
References


Clone A1 Consensus Sequence

GCGTGGTGGAATCCCACTCCTCCTATGGGATTATATCAGTAAAAAGGAGGGAGTAGTCAGGACAATAGCTGCGAAAGAAAATGCAACCCCCTAGTCCTTACAGTT
CACCAGAAGGAAAGACAAAGCTTCTTGGGACGGACCTAAGATGTGGGAGATT
ACGACTATACCCGTACGGATAGTACCCTGCTACCTTTATATTACGGTGTCGGGC
AGGTATCAAGGATACGCGCTACGCTCCGGAATGGGACCCACACTAGTCTTACC
TGATCAAGAAAACCCCATCCCGACACATCTCAACAGAGTCCAAAGTGGCAACC
CAGAGGCCAACACGAATCAAGGATACCCCAAGGCTGCTGTTGCCCGCCACCA
TGAGTCCCAAACGGATTGGGACCGGAGATAGGTTAATAAATTTAGTGCAAGGG
GACATACCTAGCTTAATAAATGCACCAGGACCCCAACAAAACTAAAGACTGTTCC
CTGCCTATCGGGTTAGTATCGGGGAAGGACTGTGCATAGGGACTGTTCC
CTAAGAACCACCAGGC

Clone A2 Consensus Sequence

GCGTGGTGGAATCCCACTCCTCCTATGGGATTATATCAGTAAAAAGGAGGGAGTAGTCAGGACAATAGCTGCGAAAGAAAATGCAACCCCCTAGTCCTTACAGTT
CACCAGAAGGAAAGACAAAGCTTCTTGGGACGGACCTAAGATGTGGGAGATT
ACGACTATACCCGTACGGATAGTACCCTGCTACCTTTATATTACGGTGTCGGGC
AGGTATCAAGGATACGCGCTACGCTCCGGAATGGGACCCACACTAGTCTTACC
TGATCAAGAAAACCCCATCCCGACACATCTCAACAGAGTCCAAAGTGGCAACC
CAGAGGCCAACACGAATCAAGGATACCCCAAGGCTGCTGTTGCCCGCCACCA
TGAGTCCCAAACGGATTGGGACCGGAGATAGGTTAATAAATTTAGTGCAAGGG
GACATACCTAGCTTAATAAATGCACCAGGACCCCAACAAAACTAAAGACTGTTCC
CTGCCTATCGGGTTAGTATCGGGGAAGGACTGTGCATAGGGACTGTTCC
CTAAGAACCACCAGGC

Clone A3 Consensus Sequence

AATCCCACCTCCTCATTGGGATTATTATCAGTAAAAAGGAGGGAGTAGTCAGGACAATAGCTGCGAAAGAAAATGCAACCCCCTAGTCCTTACAGTT
CACCAGAAGGAAAGACAAAGCTTCTTGGGACGGACCTAAGATGTGGGAGATT
ACGACTATACCCGTACGGATAGTACCCTGCTACCTTTATATTACGGTGTCGGGC
AGGTATCAAGGATACGCGCTACGCTCCGGAATGGGACCCACACTAGTCTTACC
TGATCAAGAAAACCCCATCCCGACACATCTCAACAGAGTCCAAAGTGGCAACC
CAGAGGCCAACACGAATCAAGGATACCCCAAGGCTGCTGTTGCCCGCCACCA
TGAGTCCCAAACGGATTGGGACCGGAGATAGGTTAATAAATTTAGTGCAAGGG
GACATACCTAGCTTAATAAATGCACCAGGACCCCAACAAAACTAAAGACTGTTCC
CTGCCTATCGGGTTAGTATCGGGGAAGGACTGTGCATAGGGACTGTTCC
CTAAGAACCACCAGGC
Clone A4 Consensus Sequence

CCTTACTATAGGGCAGCGTGTTGGAATCCCACTCCTCATGGGATTATATTC
CAGTAAAAAGAGGGAGTGTAGTCAGGACAATAGCTCGAAGAAAATGCAACC
CCCTAGCTTAACAGTGCCAAGGAGACCAGCAGCCTTTGGAGGACCACC
TAAGATGGGGATTACGTACTATACCGGATTAGATGCCCTTGCACTTTAT
TCACGCGTGTCGCGGACAGGTACAGGACATCAACGCCTCGGAAATGGGACC
AACTAGCTAGTTAGCTACCCGCAAGCGAAAAGATAGTACCAAGCATGGCC
TGAATGAGCTGCAATCCGCAACCACAAGGAGAGGACCTTACTTAGGATAGGTA
AAATTTAGTGAAGGGACATACCTAGGCTAAATGCAACCGACCAAAAAACAA
ACTAAAGACTGTGTGCTCTGGATATCTTGACCCACCACCACATTTGAGGAAG
AAGTTTGGCATTTATCTAGGATTTCTGAGGAGAGGACTGTGCAAGGG

Clone A5 Consensus Sequence

GCGTGTTGGAATCCCACTCCTCATGGGATTATATTCACAGTAAAAAGAGGGA
GTAGTCAGGACATAGCTCGAAGAAAATGCAACCCCATGCTTTACAGTT
CACCAAGAGGGAAAGAACACGCCCCCATTGCCCAAAGAGGACCTATGCTTTTAC
ACGACTATACGGTACCGGATAGCACCCTGTCACTTTATTCAGGTGTTCGCCGC
AGGTACGCTACGAGCCGCGCTTGCCAAGGACAAGGAAAGGACCAGCCTGGGG
AGGTTTTAGTCAATCCGCAACCACAAGGAGAGGACCTTACTAGGATAGGTA
AAATTTAGTGAAGGGACATACCTAGGCTAAATGCAACCGACCAAAAAACAA
ACTAAAGACTGTGTGCTCTGGATATCTTGACCCACCACCACATTTGAGGAAG
AAGTTTGGCATTTATCTAGGATTTCTGAGGAGAGGACTGTGCAAGGG

Clone B1 Consensus Sequence

CCTCCATGGGAATATGATGAGCTCGACCTTGTGCAAGAAAACCTGGAAG
CTCATTATAGCCAAAAAGAGGTACTCTGGGAGTACCAAGGAAAAACTATAAT
GCCCTGAAAAATACGCAAAGAGGTAGTTAGCAGATCAGGATCAGGACTGTGCTTCT
CTCAGTGCTAGAAATAGAAATCTTACTAGAAAGGAAAACCTGGAAG

23
ACCTCCCATACAGAGACTTACACCTCCGGCAAGTAACAGAGAGCTGCCGGGC
ATGTGCTCAAATATCAACGCGAGGGAAGATAAAGTTTGAGCTGACGTAAGAGCC
CGAGGACCGGCCCTGGAATACATTGGGAAGTAGACTTTACTGAAATCAAGC
CAGGAATGTATGGATATAATACCTCTTGTGGTCTCAGATAGATACCTTCTCTGGC
TGGGCCGAAGTTACCACCACCCACAAATAGAAAAACAGAAAAGCCTTGGCAAA
AACCTTTAGAAGAATAATGGACATATAGTTGAATTTTACCTCTTAATATTAAAGCTAG
AAGCTGACGCTGACGATATGGGTGCTACCGGTTACCCCAATTCAGCAGG
CTGTTTCGCCCCCTGGAGACCATATTGCCAAGACCGAGAGACTGTGGCTAGG
ACCCACCTAGTGGGACCTCAATGATTCTGGTGAGGTCTTTATATTAAACATAGACGT
GGAAATTGGCCATTCTAATGCGCCACCAAATATAAATGTAACCTGGTTAGA
ACCAAATGTAACAAACCACACCAAGCTGTCGACCTCCTTAGCTGAAACAGAGCC
TCAACCCCTTCAAAGCTGAGACGGCTCTGCTGATATCTCTAGCTTCCTACTT
CCCTGACCGGCAGTGTTGCACTGTATGCTATGACTGCTGATGCTGCAATGGGACCAA
ACCTAGTCTTACCTCCGCGACAGGTGAGATCTGCTGCTCTGCTGCTCAGGCTAG
ACCCACCTTCTTTTACGTCCCTGCCCAGGGCATGCCCCCGGAATGGGGCCAAAG
GGAACACATTGTGGAGGGGCCAACAGATGGTTTTGTGGCGCTTTGGGGATGTA
AACACAGGAAAGCTGTCGACCTCCTAGCTGCAATGACCTTCACCCCAATGCTGAA
ATCTCCCTCATAAGATGGGTGAAAACATCGAATAGAGAAAGGACAGACAA
AGATGTGGGGATTACGACTGTACCGTACCGGATATGACCCTGTCACTTTATTC
ACGGTGCTCCCGCAGGTATCAACATACATCCGCTCAGGCAATTGGAACCA
ACCTGCTTTACCTGATCAAGAACCCCATCCTCCGCAATCTAACAGAGTC
CAAAGTGCCGACCACAGATGGCTGCTGCTGCCAGCCTGCTGCTCAGGCAATTG
GAAGGATGTTAAATATAATTGAAAGTGCAACGATCATTAGCTGCTTAAAATGCAACCA
ACCCCAACAAAATAATACAGCGTTGCTCCCTAGGCTTGTTTGCAACCTCCTAT
TACGAAAGGTTTGGCAATCATAGGTAACTACAGCAACCAAAAAACAAAACACCC
CATCTGCTCTATCTACCTCCGCAATGAAAATCAGTTATCTGCTGATCGG
CAAGGACTGTCGACAGGGACTGTTGCTCTAAGAGCCAACCACGG

Clone B2 Consensus Sequence

CTATTGGGCCACGCGTGTCGACCGCCCGGGGCTGTCCTCCATGGGAATATGA
TGACAGTTTAGAGACCTGTCGACAGAAACTAAAACAGCTTATGAGCCAAACA
Clone B3 Consensus Sequence
Clone B4 Consensus Sequence

GAATGTATGGATATAATACCTCTTTGTTGGTCATAGACACCTTCTCTGGCTGG
GCCGAAGCTTACCCCGCCAAACATGIAACAGCAAAGTGTTGGCAAGAAAC
TCTTGAAGAGATTTTTCCCTGTTACGGGATCCCTCAAGTATTGGGTTTCAGAT
AATGACCACGCCTTTATCTCCAAGTGAAGTCAAGCTGTCCTGAGGCCACCTACTGGG
GATTAATTGGAATAGACTACCTTCGATACCAGCACCACCAAGTTTCAAGTCAGGTA
GAAAGAAATGATAAGTACATTAATAGGGAGACTTTAACTAAATTAAACGGCTAGAA
CTGGCTCTAAGGATGGGTGCTCCTCCTGCGCCCTGTTTATACCGGGTGACGT
AACACGCGATCCACGCGGTTAATCCTTGTGAAATCTCTGACGAGACCAAGC
CCACACCTATGCTGCACCTTCTTTTGATGCTGACATCTCTCTGTTTCGCTACCTCC
CCACCATGCGGCCATTTACGCGCCCCCTCGAAGTTAGGTCAAGGAAGGATCAA
GAGACCTGCTAGCGCAAGCCTAACCAGAAAAGTCAAGAAACCCCTGTTGCTGCT
CACCCCTTCAAAACCAGAGACCTCCTGCTGTTGGGAGACATCAAACCAAGT
GCTGCAAGCGCCAACCACCAATCGAGACCCACCGACCGTACAGACG
ATCCATCGAGATGGAAGTGCTCAACCCCGACCCAAAAAACCCTCTCATAAGATAGACT
TTCTCGTGAGCCCTAAATGATCTCTGCTGTGGGTCTTATTAAAGTACAGCTTTGG
AATGAGGCAATCTATTGACCAACCCAAATATATAAATTGAACTTGGAATAAACC
AATGTAACAAAATATGCTGATTAAAAACTGACAGATAGAAAAGAAACAGCAACAGA
CATACCCCTTTTACGTCGCTGCCAGGGCATGCCCGGAAATGAGGGCCAAAGG
AACAAATTTGGGAGGGGCAACAAAGATGGGTGTTCTGCGCTTTGGGGGTGATGGA
ACCAGAGAAGCGTTGAGAATCCACCTCCTCTACAGATTATATACAG
TAAAAGAGAGGATAGTACAGGAAATAGTGCCAAGAAATATGCAACCC
TAGCTCTACACAGTACCAGAACGGGAAGGACAAGCTCCTTGAGGACGACCTAA
GATGTGGGAATTACAGTATACCCAGGATAGTACACCCCTGCTACCTATTACA
CGCTGTCCGCGCAAGGTATACCACTACGCGCTTCCAGGCAATGGAACCA
AACCTAGTCTACTGATCGAAAGACCACCCTCAGCAATCCTAAACAGAGT
CCAAATGGCCACACCGAAGCGGGCAACAAAGAATACAGATACCCAAGGGTCTGG
TGCCACCGGCAACCAGAGTCCCAAAACGGAATGGGAGAAGATGATGTTAATA
AATTTGATGCGAGGACATACCAGCCTGAATGCGACCACCCAAAAA
CTAAAGACTGTGGCTCCTGCTTCTGATCGGACACCCTATTAGAAGAGATT
GCAATCTTAGGATACAGCAACAAACAAACAAACCCCTCATTCTCCCATC
TACTCCGCCAATAAATCAACTATATCTGAGGATACGGGGCACAAGGACTGTGCA
TAGGGAGTGTCTCTAAGGACCCACCAG
TACTATAGGGCACGCGCTGGTCGACGCGCCGGGTTTGTGCTCCTCATGGGAAATA
TGATGACAGTGTATTTAGACCTTGCTGCGAGAATTTGGAACCTTGCATTATGGACCC
AAAAAGAGGTACCTGGGAGTACCAAGGGAAAACTATAATGGCTGAAATA
CGCAAAAGAGTTGATTAGCCACTTGCATAAGTTACACACCTCACTGTTGAAG
AAAAATGAAAAGTATAGAATGAGAAAGAGAAGAACTTTGTTTACCTTCTGTTGCGCCGAGC
TTACCCCGCAAACATGAAAAAGCAGAAAAAGTGTGCTGCAAGAAAACTTTAGAGAA
GAGATTTTCTCCTCAGGGAATCCTCAGTTATGTTGTTGAGATACATATGG
AGTATAAATACCTTCTGTTGACATTTACATGGCAATTCCTCCCTCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT
GCAGGCAACATTTGCGGCTTGCAGTACTGCAAGAGATGGAAGAGATCCAAGAGACT
CTAGCGACACACGAAAAAGAAGCCGCTGACGCCAAGGGTTGCTCATCACCCT
CTAGCCGAGCCTACCGAGAAGGCCTGACGCCAAGGGTTGCTCATCACCCT
TCAAACCAGGACATCTCCTGCTGTTGAGATACATGGCAATTCCTCCCTCTCTCTTCTGACCACT
TGAAATTTACATTGTGACATACCAGACCCCCAAAAGTTCAGGTTACGAGAAAAGAAT
GAAATAGATCTTAATAGAGACCTCTTAAGTAAAATAGAGCTGAAACTAGGTCTCTT
AAGGATTGTTGTCCTTCCCTGAGCTGTTTTATAACCCGTTGACATGAACAGCC
AGTGTCTGACAAGTTTCTCTCTTCTTCTGTTGACACTCTCTCTTCTCAGATCT