

Article

# A Novel Retrovirus (Gunnison's Prairie Dog Retrovirus) Associated With Thymic Lymphoma in Gunnison's Prairie Dogs in Colorado, USA

Molly D. Butler <sup>1</sup>, Karen Griffin <sup>2</sup>, Connie D. Brewster <sup>1</sup> , Marylee L. Kapuscinski <sup>1</sup>, Mark D. Stenglein <sup>1</sup> , Daniel W. Tripp <sup>2</sup>, Sandra L. Quackenbush <sup>1,\*</sup> and Karen A. Fox <sup>2,\*</sup>

<sup>1</sup> Department of Microbiology, Immunology and Pathology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO 80523, USA; molly.butler@colostate.edu (M.D.B.); conniebrew@gmail.com (C.D.B.); mllayton@rams.colostate.edu (M.L.K.); mark.stenglein@colostate.edu (M.D.S.)

<sup>2</sup> Colorado Division of Parks and Wildlife, Wildlife Health Laboratory, Fort Collins, CO 80521, USA; karen.griffin@state.co.us (K.G.); dan.tripp@state.co.us (D.W.T.)

\* Correspondence: sandra.quackenbush@colostate.edu (S.L.Q.); karen.fox@state.co.us (K.A.F.); Tel.: 1-970-491-3545 (S.L.Q.); Tel: 1-970-472-4318 (K.A.F.)

Received: 8 May 2020; Accepted: 31 May 2020; Published: 2 June 2020



**Abstract:** As part of research and wildlife disease surveillance efforts, we performed necropsy examinations of 125 free-ranging ( $n = 114$ ) and captive ( $n = 11$ ) prairie dogs in Colorado from 2009 to 2017. From these cases, we identified three cases of thymic lymphoma in free-ranging Gunnison's prairie dogs (*Cynomys gunnisoni*), and we identified a novel retroviral sequence associated with these tumors. The viral sequence is 7700 nucleotides in length and exhibits a genetic organization that is consistent with the characteristics of a type D betaretrovirus. The proposed name of this virus is Gunnison's prairie dog retrovirus (GPDRV). We screened all 125 prairie dogs for the presence of GPDRV using PCR with envelope-specific primers and DNA extracted from spleen samples. Samples were from Gunnison's prairie dogs ( $n = 59$ ), black-tailed prairie dogs (*Cynomys ludovicianus*) ( $n = 40$ ), and white-tailed prairie dogs (*Cynomys leucurus*) ( $n = 26$ ). We identified GPDRV in a total of 7/125 (5.6%) samples including all three of the prairie dogs with thymic lymphoma, as well as spleen from an additional four Gunnison's prairie dogs with no tumors recognized at necropsy. None of the GPDRV-negative Gunnison's prairie dogs had thymic lymphomas. We also identified a related, apparently endogenous retroviral sequence in all prairie dog samples. These results suggest that GPDRV infection may lead to development of thymic lymphoma in Gunnison's prairie dogs.

**Keywords:** *Cynomys gunnisoni*; endogenous retrovirus; prairie dog; retrovirus; thymic lymphoma; tumor; *Yersinia pestis*

## 1. Introduction

Colorado is home to three of the five free-ranging prairie dog species native to North America: the white-tailed prairie dog (*Cynomys leucurus*), the black-tailed prairie dog (*Cynomys ludovicianus*), and the Gunnison's prairie dog (*Cynomys gunnisoni*). Prairie dogs are a member of the *Sciuridae* family, along with squirrels and other burrowing rodents. Prairie dogs are considered "keystone" species in that prairie dogs and their burrows provide prey and habitat for a number of wild birds and mammals, they support diverse plant and pollinator communities and the health of the local ecosystem often depends on the health of prairie dog populations [1,2]. Disease threats from plague (caused by *Yersinia pestis*) are well understood in prairie dogs [3,4], which are highly susceptible to

fatal infections from *Y. pestis*. However, the roles of other infectious agents in the health of free-ranging prairie dog colonies have not been extensively considered.

Much of the available information about prairie dog health, including the incidence of cancer, comes from studies in prairie dogs that are kept in captivity either as pets or for research. The most commonly reported tumors of captive prairie dogs are elodontoma and hepatocellular adenocarcinoma, of which hepatocellular carcinoma may be associated with infection by a hepadnavirus [5,6]. Other tumors are less commonly encountered in captive prairie dogs but do include lymphoma and thymoma [5,7,8]. An infectious cause for lymphoid tumors in prairie dogs has not been investigated, although retroviruses are a well-known cause of several human and veterinary lymphoid cancers [9–15]. In some cases, such as infection with feline leukemia virus and feline immunodeficiency virus, possible sequela of retroviral infection also include immune suppression [16–20].

During the course of wildlife disease research and surveillance activities in Colorado, USA, we identified three free-ranging Gunnison's prairie dogs with thymic lymphoma, and a novel type D betaretrovirus associated with these tumors. Based on the known sensitivity of prairie dogs to sylvatic plague, we also considered retroviral infections as a possible source of immune suppression that could contribute to mortality.

## 2. Materials and Methods

### 2.1. Necropsy, Histopathology, and Immunohistochemistry

From 2009 to 2017, we necropsied 125 free-ranging ( $n = 114$ ) and captive ( $n = 11$ ) prairie dogs from Colorado. Species examined included Gunnison's ( $n = 59$ ), black-tailed ( $n = 40$ ), and white-tailed ( $n = 26$ ) prairie dogs. Animals examined were either found dead ( $n = 118$ ), died during processing ( $n = 3$ ), or were euthanized due to disease concerns ( $n = 4$ ). Research methods included trapping and brief anesthesia with isoflurane gas [21] and were approved (06/06/2013) by the Colorado Parks and Wildlife Animal Care and Use Committee #06-2013.

Gross necropsy was performed for all 125 prairie dogs to determine cause of death, and we pursued histopathology if cause of death was not apparent from gross necropsy and tissues were suitable. Prior to necropsy, the carcasses were either frozen at  $-20\text{ }^{\circ}\text{C}$  to preserve the carcass and to kill fleas, or fresh carcasses were treated with insecticide (Deltamethrin/DeltaDust, Bayer Environmental Science, Cary, NC, USA) prior to necropsy to kill fleas without the need for freezing. Necropsies were conducted in a biological safety cabinet (NU-S813-400, Nuair, Plymouth, MN, USA) with additional personal protective equipment in accordance with BSL-2 biosafety practices. After necropsy, carcasses were frozen at  $-20\text{ }^{\circ}\text{C}$  until *Y. pestis* PCR results were obtained, and any carcasses with tissues confirmed positive for *Y. pestis* or *Francisella tularensis* were disposed of by chemical digestion. For histopathology, tissues were fixed in 10% neutral buffered formalin, embedded in paraffin wax, sectioned by microtome to approximately 8 micrometers, affixed to glass slides, and stained with hematoxylin and eosin.

We used immunohistochemistry to identify T-lymphocytes (CD-3 (LN10, Leica Biosystems, Buffalo Grove, IL, USA)), B-lymphocytes (PAX-5 (1EW, Leica Biosystems, Buffalo Grove, IL, USA)), and epithelial cells/cytokeratin (MCK (AE1/AE3, Leica Biosystems, Buffalo Grove, IL, USA)) in formalin-fixed paraffin embedded tissues. The above monoclonal mouse anti-human antibodies were applied using a Leica BOND-MAX automated IHC staining platform (Leica Biosystems, Buffalo Grove, IL, USA), with chromogen Poly-AP anti-mouse (PV6110, PowerVision, Leica Biosystems, Buffalo Grove, IL, USA) used for PAX-5 and MCK, and chromogen Poly-HRP anti-mouse (PV6113, PowerVision, Leica Biosystems, Buffalo Grove, IL, USA) used for CD-3. Slides were counterstained with hematoxylin. Negative controls of duplicate tissue sections were incubated in antibody diluent and homologous nonimmune sera. Non-specific staining was not observed. To confirm efficacy in prairie dog tissues, we applied IHC stains to control tissues, including thymus, spleen, and skin from a yearling prairie dog that died from enteric disease. These control tissues demonstrated expected staining properties including robust staining of T-lymphocytes in the thymus with CD-3, scattered

staining of B-lymphocytes in the thymus with PAX-5, and staining of epithelial cords and nests (Hassall's corpuscles) in the thymus with MCK. Control prairie dog spleen demonstrated robust staining with PAX-5, highlighting follicular structure, and skin epithelium demonstrated robust staining with MCK.

## 2.2. PCR for *Yersinia Pestis* and *Francisella Tularensis*

We extracted DNA from spleen tissue of all ( $n = 125$ ) prairie dogs, under BSL-2 conditions, using a DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA). Each sample was tested for presence of *Y. pestis* by PCR using primers *caf1-F/caf1-R* (Table 1) [22] and cycling conditions as previously described [23]. Primers (50 $\mu$ M) and DNA template (50–250 ng) were added to a 0.2 mL PCR tube containing a puReTaq Ready-To-Go PCR bead (Illustra, GE Healthcare Bio-Sciences Corp, Piscataway, NJ, USA) for a final volume of 25  $\mu$ L. Cycling conditions were: 94 °C for 10 min (1 cycle), followed by 94 °C for 1 min, 55 °C for 1 min, 72 °C for 30 s (35 cycles), and 72 °C for 10 min (1 cycle). The final product was visualized on a 2% agarose gel. DNA extracted from spleen (see above) was also tested for *F. tularensis* using primers P2/P3 (Table 1) and cycling conditions as previously described [24]. Primers (20  $\mu$ M) and DNA template (50–250 ng) were added to a 0.2 mL PCR tube containing a puReTaq Ready-To-Go PCR bead (Illustra, GE Healthcare Bio-Sciences Corp, Piscataway, NJ, USA) for a final volume of 25  $\mu$ L. Cycling conditions were: 97 °C for 10 min (1 cycle), followed by 94 °C for 1 min, 55 °C for 1 min, 75 °C for 1 min (35 cycles), and 75 °C for 10 min (1 cycle). The final product was visualized on a 2% agarose gel.

**Table 1.** PCR primers.

Primer Designation	Primer Sequence
<i>caf1-F</i>	5' ATA CTG CAG ATG AAA AAA ATC AGT TCC 3'
<i>caf1-R</i>	5' ATA AAG CTT TTA TTG GTT AGA TAC GGT 3'
P2	5' TAG GAT CCC ATT AGC TGT CCA CTT ACC 3'
P3	5' GGA ATT CGT TAG GTG GCT CTG ATG AT 3'
YMDD	5' ATC AGA TCC TAC TAA CDR TCR TCC ATR TA 3'
LPQG	5' TAC CAG TGG AAT GTT CTA CCN 3'
LTR-F	5' GAC CGT GAC TTG TTT ATC TAA CCA CAA 3'
Env-7073	5' AGA CTG CAA TCT TTG GTA ATG AAC CTG 3'
SP3	5' AAG GTT CTT CAT CCA GGA GGT ATA TCT C 3'
SP4	5' TTT GGA CTT CAA CCA TGG GGA AAA TAC 3'
SP5	5' TTC CGT ACT TAC CCC TTC TTT CCG AT 3'
LTR-R	5'CCA AGG TTC TTC ATC CAG GAG GTA TAT 3'
494F	5' AAG GAT GTG AAG GAA CTA TAC AGC CAT 3'
591R	5' TCG GGG TGA ATT GGA ATT GAA AAG AAA 3'
393F	5' ATGCTGACTGGGATATGGTCAAAAATG 3'
1961R	5' GTATCGGGTTTCCTTGGACATCAAATT 3'
5729F	5' GGATGGCAAATCGTATTATCAGGCTAC 3'
5840R	5' TAGAGTCCCCTGAGGTACCTAAGAT 3'

## 2.3. RNA Extraction and RT-PCR

Total RNA was extracted from fresh-frozen spleen (case nos. 11-1310, 14-1342, 15-1406) and thymic tumor (case nos. 14-1342, 15-1406) tissue samples using TRIzol (ThermoFisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. One microgram of DNase (Ambion Turbo DNase; ThermoFisher Scientific, Waltham, MA, USA) RNA was reverse transcribed into cDNA using SuperScript III (ThermoFisher, Scientific, Waltham, MA, USA) according to the manufacturer's protocol. A 3:1 mix of random hexamers and oligo-dT or the degenerate primer, YMDD, was used as the reverse primer (Table 1). PCR was performed using *Taq* DNA polymerase (ThermoScientific, Waltham, MA) with the degenerate primers LPQG and YMDD (Table 1). PCR reactions contained 2  $\mu$ L cDNA in reaction buffer comprised of 200 nM of each primer, 2 mM MgCl<sub>2</sub>, 0.2 mM dNTPs and 0.5 units of

Fermentas *Taq* DNA polymerase (ThermoFisher Scientific, Waltham, MA, USA) in a total volume of 50  $\mu$ L. The cycling conditions were as follows for cycles 1–10: 94 °C for 1 min, 37 °C for 2 min, and 72 °C for 3 min. The conditions for cycles 11–40 were: 94 °C for 30 s, 55 °C for 1 min, and 72 °C for 1 min [25]. The PCR product was purified (QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA)), cloned into the pCR2.1 TOPO vector (ThermoFisher Scientific, Waltham, MA, USA) and submitted for sequencing.

#### 2.4. Library Preparation, NextGen Sequencing, Genome Assembly and Analysis

The KAPA Biosystems RNA HyperPrep Kit (Roche, Pleasanton, CA, USA) was used to prepare sequencing libraries from 100 ng total RNA isolated from the lymphoma and spleen from two Gunnison's prairie dogs (14-1342, 15-1406) with tumors, the spleen from two Gunnison's prairie dogs (15-656, 15-671) without tumor, and the spleen from a Black-tailed prairie dog (14-1382) without tumor according to the manufacturer's protocol using half-scale reactions without fragmentation. Pooled libraries were length-selected for 300–500-bp fragments using a BluePippin 2% cassette (Sage Biosciences, Beverly, MA, USA). Length-selected libraries were cleaned using a 1:1.4 ratio of solid phase reversible immobilization [26] beads (Kapa Biosystems, Roche, Pleasanton, CA, USA). Individual libraries were then pooled for sequencing. The library pool was diluted to 4 nM based on fluorometric DNA quantification (Qubit High Sensitivity DNA Assay; ThermoFisher Scientific, Waltham, MA, USA) and quantified by qPCR using the KAPA Library Quantification Kit (Kapa Biosystems, Roche, Pleasanton, CA, USA). Paired-end 2  $\times$  150 bp sequencing was performed on an Illumina NextSeq, producing an average of 5  $\times$  10<sup>6</sup> read pairs per dataset.

Datasets were processed as previously described, with the goal of taxonomically categorizing all non-prairie dog reads [27]. Briefly, first low quality and adapter sequences were filtered using Cutadapt v1.18 [28]. Duplicate read pairs (reads that shared >96% pairwise identity) were removed using cd-hit v4.8.1 [29]. Bowtie2 was used to remove host-derived reads by mapping to a combined index built from the *Marmota marmota* (European marmot; GCF\_001458135.1) and *Ictidomys tridecemlineatus* (thirteen-lined ground squirrel; GCF\_000236235.1) genomes and transcriptomes [30,31]. Reads with an alignment score >60 were removed. Remaining reads were assembled using the SPAdes assembler [32]. Contigs were taxonomically assigned by searching the NCBI nt database using BLASTN and then by searching the NCBI nr database using diamond [33,34]. Candidate retrovirus-derived contigs were manually inspected in Geneious Prime 2020.0.3 (<https://www.geneious.com>) and validated by re-mapping reads using bowtie2 as above. This analysis pipeline is available at [https://github.com/stenglein-lab/taxonomy\\_pipeline](https://github.com/stenglein-lab/taxonomy_pipeline). Libraries from HeLa cell total RNA and water were constructed and analyzed in parallel as positive and negative controls.

PCR and Sanger sequencing confirmation of the Gunnison's prairie dog retrovirus (GPDRV) genome assembly was performed on high-molecular weight splenic DNA from a tumor-negative but GPDRV sequence positive prairie dog using a forward primer in the LTR (LTR-F) and a reverse primer in *env* (Env-7073). PCR conditions were as follows: 98 °C for 30 s, then 30 cycles of 98 °C for 10 s, 66 °C for 20 s, and 72 °C for 4min, followed by 72 °C for 5 min. A single PCR product of the expected size (7266 bp) was cloned into Strataclone Blunt vector (Agilent Technologies, LaJolla, CA, USA) Plasmid DNA was isolated from individual colonies. Samples were confirmed positive for the insert by restriction digest and sent for Sanger sequencing. The sequence was confirmed using primer walking.

Fisher's exact test was performed as implemented in R [35].

#### 2.5. Phylogenetics

To collect relevant Pol sequences, we used two strategies. First, we queried the GPDRV Pol sequence against the NCBI nr protein database using BLASTP [36] and retrieved all aligning sequences that produced alignments with E-values lower than 10<sup>-40</sup>. We removed sequences shorter than 600 amino acids and used cd-hit to collapse sequences that shared >95% pairwise identity [28,29]. Secondly,

we collected all Pol sequences in the NCBI RefSeq protein database annotated under the family *Retroviridae* (taxid 11632). To collect Env transmembrane protein (TM) domain sequences, we collected all Env sequences in the NCBI RefSeq protein database annotated under the family *Retroviridae*. Because of the high level of sequence divergence between retrovirus Env sequences, we selected the subset of these Env TM refseqs that produced a blastp alignment with an E-value  $<10^{-40}$ .

In all of these cases, we aligned collected sequences using the MAFFT aligner v7.407 with default parameters and trimmed alignments using TrimAL v1.4.rev15 [37,38]. The best model for tree inference was selected using modeltest-ng and trees were created using RaxML-ng v. 0.9.0 with standard parameters [39,40]. Trees were visualized using the Interactive Tree Of Life (iTOL) v4 [41].

## 2.6. Integration Site Analysis

We performed an integration site analysis using Retro-X Integration Site Analysis Kit (Clontech, Mountain View, CA, USA) according to the manufacturer's instructions. Briefly, high molecular weight genomic DNA was isolated from tumor tissue and then digested with restriction enzymes *Ssp1*, *Hpa1*, or *Dra1*. Digested DNA was purified and then ligated to GenomeWalker (Clontech, Mountain View, CA, USA) adaptors using T4 DNA ligase. A primary PCR reaction was performed on the adaptor-ligated DNA using an outer, adaptor-specific forward primer AP1, and an outer, GPDRV sequence-specific reverse primer (SP3, SP4, or SP5, Table 1). A secondary, or nested PCR reaction was performed using the primary PCR reaction amplicons as template, a nested adaptor-specific primer AP2, and a nested PDRV sequence-specific reverse primer (SP3 or LTR-R, Table 1). Secondary PCR products were visualized by electrophoresis on an agarose gel to confirm a single, predominant PCR product. PCR products were cloned (Topo TA (ThermoFisher, Waltham, MA, USA)) and plasmid DNA was isolated from individual colonies. Samples were confirmed positive for inserts by restriction digest and submitted for Sanger sequencing. Sequences that partially overlapped on the 3' end with the 5' end of the PDRV sequence, but that diverged upstream were identified as integration sites.

## 2.7. PCR Screening for GPDRV and ERV-PDRV.1-Cynomys Ludovicianus

Endogenous pro/pol: Primers 494F/591R (Table 1), were used to screen DNA extracted from forty-two prairie dog spleen samples (DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA)). These samples were selected to include captive ( $n = 7$ ) and free-ranging ( $n = 35$ ) animals representing the timespan of the entire project, and originating from carcasses with minimal to mild autolysis suggesting good DNA quality. DNA template (10 ng) was added to a PCR mixture containing 5  $\mu$ L 10x *Taq* buffer with  $(\text{NH}_4)_2\text{SO}_4$ , 0.5  $\mu$ L *Taq* polymerase (Fermentas, ThermoFisher Scientific, Waltham MA, USA), 200  $\mu$ M (each) dNTPs, 500 pmol (each primer), 1 mM  $\text{MgCl}_2$ , and sterile water to a final volume of 50  $\mu$ L. Cycling conditions were: 95 °C for 3 min (1 cycle), followed by 95 °C for 30 s, 53.7 °C for 30 s, 72 °C for 30 s (30 cycles), and 72 °C for 5 min. The final product was visualized on a 1% agarose gel.

GPDRV gag gene: Primers 393F/1961R (Table 1), were used to screen DNA extracted from forty-two prairie dog spleen samples (DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA)). DNA template (10 ng) was added to a PCR mixture containing 5  $\mu$ L 10x *Taq* buffer with  $(\text{NH}_4)_2\text{SO}_4$ , 0.5  $\mu$ L *Taq* polymerase (Fermentas, ThermoFisher Scientific, Waltham, MA, USA), 200  $\mu$ M (each) dNTPs, 500 pmol (each primer), 1 mM  $\text{MgCl}_2$ , and sterile water to a final volume of 50  $\mu$ L. Cycling conditions were: 95 °C for 3 min (1 cycle), followed by 95 °C for 30 s, 51 °C for 45 s, 72 °C for 30 s (30 cycles), and 72 °C for 5 min. The final product was visualized on a 1% agarose gel.

GPDRV env gene: Primers 5729F/5840R (Table 1), were used to screen DNA extracted from prairie dog spleen samples (DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA)). DNA template (10 ng) was added to a PCR mixture containing 5  $\mu$ L 10x *Taq* buffer with  $(\text{NH}_4)_2\text{SO}_4$ , 0.5  $\mu$ L *Taq* polymerase (Fermentas, ThermoFisher Scientific, Waltham, MA, USA), 200  $\mu$ M (each) dNTPs, 500 pmol (each primer), 1 mM  $\text{MgCl}_2$ , and sterile water to a final volume of 50  $\mu$ L. Cycling conditions were: 95 °C for

3 min (1 cycle), followed by 95 °C for 30 s, 54.3 °C for 30 s, 72 °C for 30 s (30 cycles), and 72 °C for 5 min. The final product was visualized on a 1% agarose gel.

Primers 5729F/5840R (Table 1) were also used to screen DNA extracted from an additional eighty-three prairie dog spleen samples (DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA)). Primers (0.5 µM) and DNA template (50–250 ng) were added to a 0.2 mL PCR tube containing a puReTaq Ready-To-Go PCR bead (Illustra, GE Healthcare Bio-Sciences Corp, Piscataway, NJ, USA) for a final volume of 25 µL. Cycling conditions were: 95 °C for 3 min (1 cycle), followed by 95 °C for 30 s, 54.3 °C for 30 s, 72 °C for 30 s (30 cycles), and 72 °C for 5 min (1 cycle). The final product was visualized on a 2% agarose gel.

### 2.8. Deposition of Sequences and of Expression Data

Sequences have been deposited in GenBank under accession numbers MT361316, MT361317, and MT361318. The data have been deposited with links to BioProject accession number PRJNA631279 in the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>).

## 3. Results

### 3.1. Thymic Lymphoma in Gunnison's Prairie Dogs

Causes of death in prairie dogs included bacteremia, trauma, capture-related factors, environmental factors, intraspecific aggression, euthanasia for disease concern, and thymic lymphoma (Table 2). Plague, caused by *Y. pestis*, accounted for nearly all (49/54; 91%) of the cases of bacterial disease. Two prairie dogs had severe lesions of tularemia, caused by *F. tularensis*. Three Gunnison's prairie dogs were affected by thymic lymphoma. No other tumors were observed in any of the other 122 prairie dogs examined. The three cases of thymic lymphoma are further described below.

**Table 2.** Causes of death in 125 prairie dogs in Colorado from 2009 to 2018.

Cause of Death	#	G <sup>a</sup>	BT <sup>b</sup>	WT <sup>c</sup>	TL <sup>d</sup>	GPDRV <sup>e</sup>
Thymic lymphoma only	2	2	0	0	2	2
Bacterial	54	21	23	10	0	3
Plague	49	20	19	10	0	2
Tularemia	2	1	1	0	0	1
Other bacteremia	3	0	3	0	0	0
Trauma	21	14	3	4	0	0
Capture related	10	9	1	0	1	2
Environmental	8	3	5	0	0	0
Intraspecific aggression	6	2	1	3	0	0
Euthanized for disease concern	4	1	3	0	0	0
Undetermined	20	7	4	9	0	0

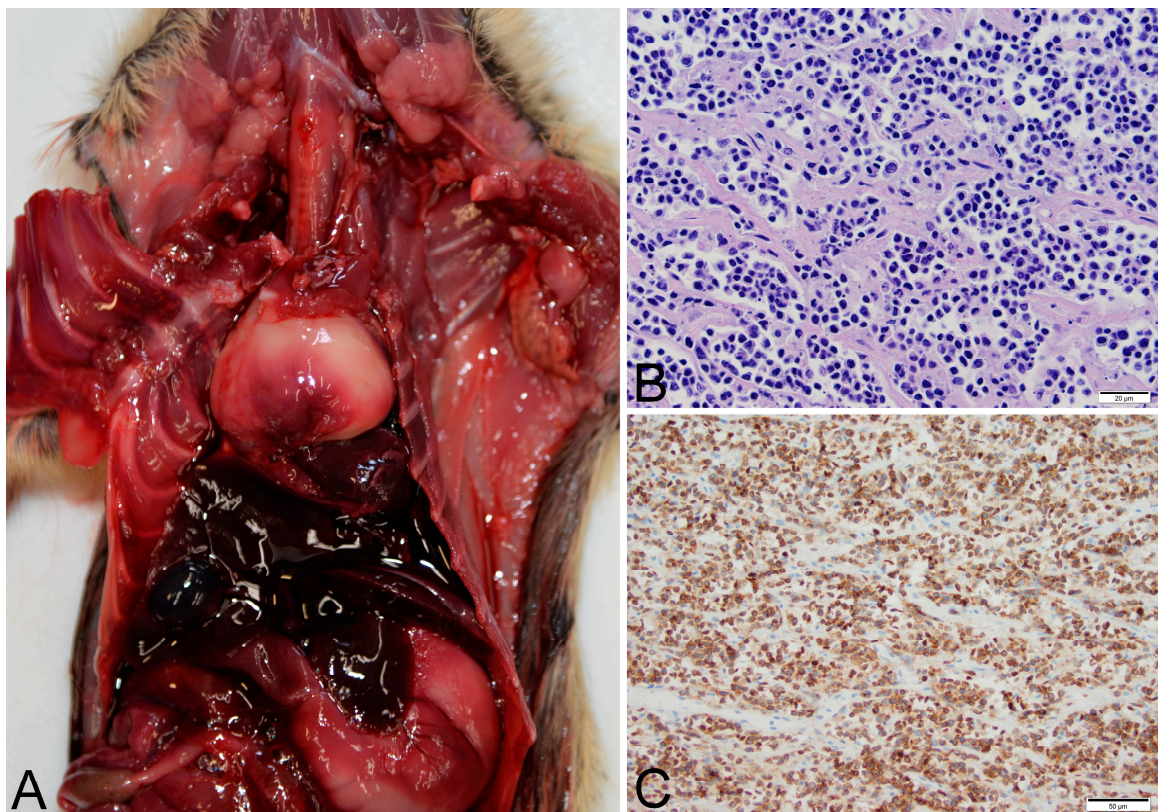
<sup>a</sup> G Gunnison's prairie dog. <sup>b</sup> BT black-tailed prairie dog. <sup>c</sup> WT white-tailed prairie dog. <sup>d</sup> TL thymic lymphoma.

<sup>e</sup> GPDRV Gunnison's prairie dog retrovirus.

Thymic lymphoma case number 11-1310 was a free-ranging yearling female Gunnison's prairie dog found dead lying half-way out of a burrow with no evidence of trauma. The chest was filled with cloudy fluid and few fibrin strands. The lungs were mottled and the heart was dilated. The thymus was enlarged to approximately 10× normal size by a soft, white mass with a mottled appearance suggesting multifocal hemorrhage and necrosis. Histopathology was complicated by freeze/thaw artifacts but, demonstrated uniform infiltrates of neoplastic lymphocytes determined to be T-cells by immunohistochemistry for CD-3. The tumor did not contain a significant population of B-cells or epithelial cells as determined by immunohistochemistry for Pax5 and MCK. Neoplastic cells invaded beyond the tissue capsule of the mass and infiltrated the surrounding adipose tissue. The tumor was

diagnosed as T-cell lymphoma. Metastasis was not observed. The spleen and liver were PCR negative for *Y. pestis* and *F. tularensis*.

Thymic lymphoma case number 14-1342 was a free-ranging, lactating adult female Gunnison's prairie dog trapped as part of research and management activities. The prairie dog was observed to have slightly labored breathing when found in the trap but did not raise concern for illness. The same prairie dog had been captured approximately one-year prior without complications. While anesthetized, the prairie dog was noted to have stopped breathing. Oxygen was administered but the prairie dog never recovered from anesthesia. At necropsy, erythema of the skin was observed on the vulva and lower limbs. An approximately 1 cm diameter granuloma was present within the mesenteric adipose tissue of the abdomen. Firmly adhered to the base of the trachea was an approximately 3 cm diameter mass that displaced the heart caudally (Figure 1). The mass was white and firm, with a mottled appearance on cut surface suggesting hemorrhage and necrosis. Histopathology was complicated by freeze/thaw artifacts, but demonstrated uniform infiltrates of neoplastic lymphocytes (Figure 1) determined to be T-cells by immunohistochemistry for CD-3 (Figure 1). The tumor did not contain a significant population of B-cells or epithelial cells as determined by immunohistochemistry for Pax5 and MCK. The tumor was diagnosed as T-cell lymphoma. Metastasis was not observed. PCR of liver and spleen were negative for *Y. pestis* and *F. tularensis*.



**Figure 1.** Thymic lymphoma identified in a Gunnison's prairie dog (Case 14-1342). (A) An approximately 3 cm diameter, soft, white mass displaces the heart caudally. (B) Histologic findings included uniform infiltrates of neoplastic lymphocytes (hematoxylin and eosin). (C) Immunohistochemical staining with anti-CD3 identified neoplastic cells as T-lymphocytes.

Thymic lymphoma case number 15-1406 was a free-ranging, approximately 1–2 year old female Gunnison's prairie dog found dead in a burrow with no evidence of trauma. Body condition was good, with plentiful fat stores in the abdomen. The chest contained an approximately 4 cm diameter mass, with red-tinged fluid filling the chest cavity. The mass was white, soft, and contained multifocal

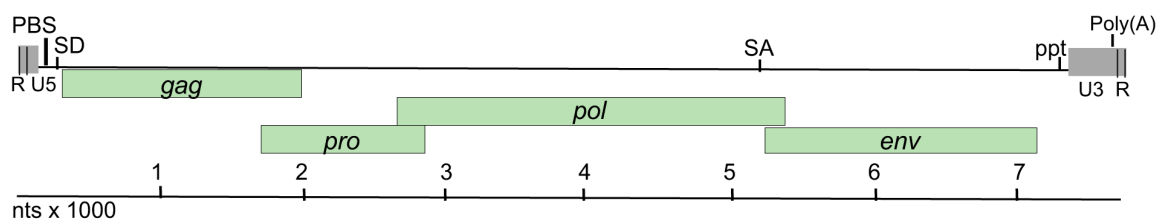
hemorrhages on cut surface. The spleen was moderately enlarged, and the inguinal lymph nodes were dark red. Histopathology was complicated by freeze/thaw artifacts, but the thymic mass demonstrated uniform infiltrates of neoplastic lymphocytes determined to be T-cells by immunohistochemistry for CD-3. The tumor did not contain a significant population of B-cells or epithelial cells as determined by immunohistochemistry for Pax5 and MCK. The tumor was diagnosed as T-cell lymphoma. Metastasis was not observed. PCR of liver and spleen were negative for *Y. pestis* and *F. tularensis*.

### 3.2. Identification of Prairie Dog Retroviral Sequences Associated with Thymic Lymphoma

Initial investigation for a possible retroviral etiology involved use of reverse transcription polymerase chain reaction (RT-PCR) with degenerate retrovirus primers (LPQG and YMDD) targeting a well-conserved region in the reverse transcriptase gene [25]. Sequencing of the PCR products from tumors 14-1342 and 15-1406 identified two unique sequences with homology to known retroviruses.

To obtain additional sequences of these two potential viruses we utilized a meta-genomics approach. Total RNA was isolated from the lymphoma and spleen from two Gunnison's prairie dogs (14-1342, 15-1406) with tumors, the spleen from two Gunnison's prairie dogs (15-656, 15-671) without tumor, and the spleen from a black-tailed prairie dog (14-1382) without tumor. An individual library was prepared from each sample and sequenced on an Illumina NextSeq. Sequencing produced an average of  $5 \times 10^6 \times 2 \times 150$  read pairs per sample. Following removal of low quality and adapter sequences, duplicate reads, and host-derived reads, an average of  $2.4 \times 10^4$  read pairs remained in each dataset (0.5% of starting datasets). These remaining reads were assembled and the resulting contigs were taxonomically classified.

Retroviral sequences were identified in the lymphoma and spleen from the two tumor-positive animals, but not in the spleen from the three tumor negative animals. A complete sequence of 7700 nucleotides (nt) in length was assembled from both tumors. The organization of the assembled genome appeared to be similar to that of betaretroviruses with *gag*, *pro*, *pol* and *env* genes (Figure 2). The nucleotide sequences of the two assembled genomes were 99.2% identical. The two sequences differ at 57 nucleotides resulting in eight amino acid changes in Gag, two in Pro, one in Pol and five in Env. The *pro* and *pol* genes are predicted to utilize ribosomal frameshifting for expression of *gag-pro* and *gag-pro-pol* precursors. The *env* transcript is likely generated by splicing. These data represent the first report of retroviral sequences associated with lymphoma in Gunnison's prairie dogs, with the provisional name Gunnison's prairie dog retrovirus (GPDRV).



**Figure 2.** Gunnison's prairie dog retrovirus genome organization. Diagram of the assembled GPDRV genome from tumors from prairie dogs 14-1342 and 15-1406. The Gag-Pro-Pol polyprotein is predicted to be translated from the genome by ribosomal frameshifting. Location of predicted sites-PBS-primer binding site, SD-splice donor, SA-splice acceptor, ppt-polyurine tract, poly(A)-polyadenylation sequence.

### 3.3. Prairie Dog Endogenous Retroviral Sequence

A second, distinct retroviral sequence was identified and a consensus contig was assembled from the black-tailed prairie dog (14-1382) spleen sample. This sequence is 6132 nucleotides in length and includes predicted *gag*, *pro*, and *pol* genes. An *env*-coding region and LTR sequences were not definitively identified. Sequence reads from the tumors and all spleen samples (animals 14-1342, 15-1406, 15-656, 15-671 and 14-1382) align with the 14-1382 consensus contig suggesting this



sequence is likely an endogenous retrovirus, provisionally named ERV- PDRV.1-*Cynomys ludovicianus*. This sequence shares 54% nucleotide identity with the GPDRV *gag-pro-pol* sequence.

### 3.4. Features of the Gunnison's Prairie Dog Retrovirus Sequence, Predicted Proteins and Phylogenetic Analysis

#### 3.4.1. LTR and Untranslated Regions

The long terminal and the 5' and 3' untranslated regions of retroviral genomes contain regulatory sequences that are central for viral replication. The LTR of GPDRV is 403 bases in length and is bound by inverted repeat sequences CAAG (nt 94–97) and CTTG (nt 7376–7379) that are essential for integration. U3 is 308 bp, R is 15 bp and U5 is 79 bp (Figure 2). The U3 region is preceded by a polypurine tract (nt 7358–7375), the site for initiation of plus-strand synthesis of viral DNA during retroviral replication. The U3 region in the 5' LTR serves as the promoter and enhancer for transcription of viral RNA. A consensus TATA box (TATATAA) is located 29 bp upstream of the predicted transcription initiation site. Binding sites for the transcription factors, NF1, AP1, Elk1, and NF-AT are present in the U3 region. The highly conserved polyadenylation signal, AATAAA, is located in the 3' LTR at nt 7680–7686, similar in position to that of the betaretroviruses, Jaagsiekte sheep retrovirus (JSRV), and enzootic nasal tumor virus (ENTV) [42,43].

The 5' untranslated region harbors a predicted primer-binding site (PBS) with sequence complementary to the 15 bases at the 3' end of tRNA<sup>Gln</sup> that would serve as the site for reverse transcriptase to initiate minus-strand DNA synthesis. A predicted splice donor site for the generation of the subgenomic *env* transcript is located at nt 124–131 in the untranslated region between the PBS and start of *gag*.

#### 3.4.2. Gag

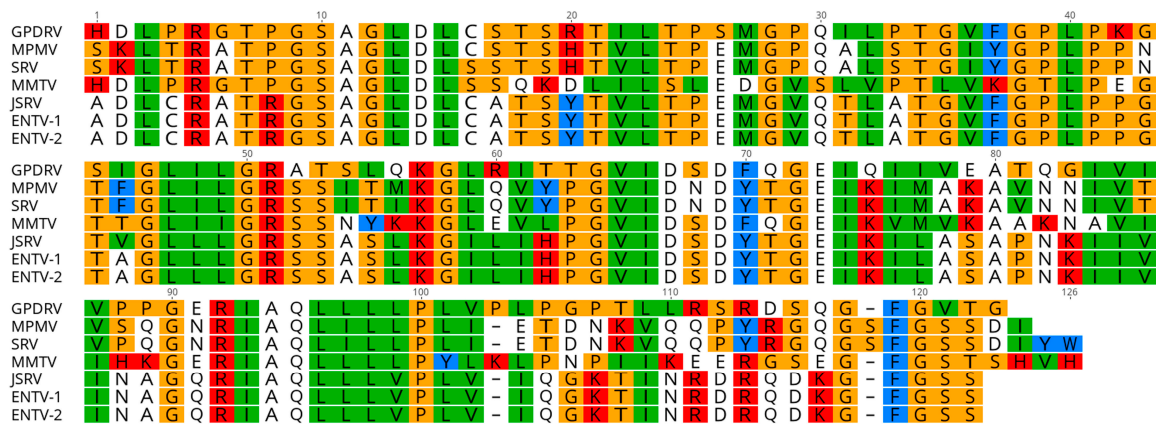
The *gag* open reading frame (nt 212–1915) is predicted to encode a 567 amino acid (aa), 62.8 kDa polyprotein. The *n* terminus of GPDRV Gag contains a consensus myristylation motif (Met-Gly) like that of many retroviruses. The GPDRV capsid protein (CA) is predicted to be a 206 aa, 22.8 kDa protein. A highly conserved major homology region (MHR), QGPSESYSDFIGRLMQSA, is located in CA. Two Cys-His motifs (Cys-X<sub>2</sub>-Cys-X<sub>4</sub>-His-X<sub>4</sub>-Cys) separated by 14 amino acids are located in the nucleocapsid protein (NC) at nt positions 1565–1607 and 1649–1690. The NC protein is predicted to be 14.3 kDa.

#### 3.4.3. Pro

The *pro* open reading frame (nt 1705–2712) is predicted to encode a 334 aa, 35.7 kDa protein expressed as a Gag-Pro fusion polypeptide generated by ribosomal frameshifting. The protein encoded by the *pro* open reading frame is comprised of two domains, similar to that of betaretroviruses: a pseudoprotease domain with dUTPase activity and the active protease (Figure 3). GPDRV dUTPase exhibits 54–57% amino acid identity with other betaretroviruses. The active protease site with a core aspartyl protease sequence, Leu-Asp-Thr-Gly, is located at amino acid 198–201 (nt 2296–2307). A glycine-rich G patch domain similar in sequence to that found in betaretroviruses is present near C-terminus of protease [44,45].

#### 3.4.4. Pol

The *pol* open reading frame (nt 2685–5375) is predicted to encode a peptide of 896 aa with a molecular mass of 101.8 kDa. Pol is predicted to be expressed as a Gag-Pro-Pol polypeptide generated by ribosomal frameshifting. The Pol polypeptide encodes reverse transcriptase and integrase activity. The conserved polymerase sequences LPQG and YMDD are located at amino acids 157 and 191 (nt 3153 and 3255), respectively. RT contains an RNase H domain with a conserved active site (DEDD). There is an *n*-terminal Zn<sup>+</sup> binding domain present in the integrase protein.

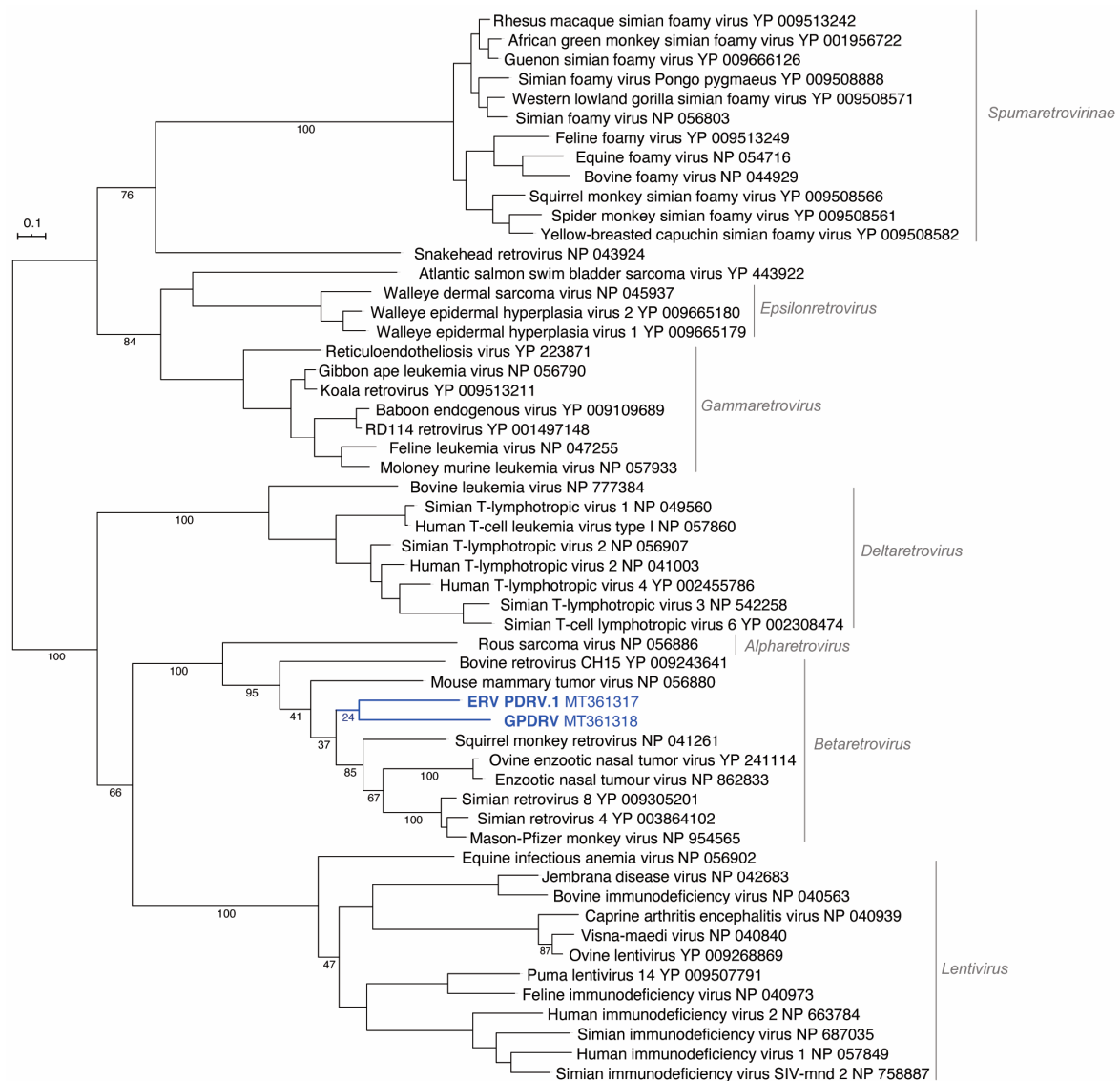


**Figure 3.** Amino acid alignment of the dUTPase domain from several betaretroviruses. Amino acid alignment was constructed using sequences GPDRV- Gunnison’s prairie dog retrovirus (this manuscript), MPMV- Mason-Pfizer monkey virus (NC\_001550), SRV (M11841), MMTV (NC\_001503), JSRV (NC\_001494), ENTV-1 (NC\_007015), and ENTV-2 (NC\_00494) with Geneious Prime 2019.2.3.

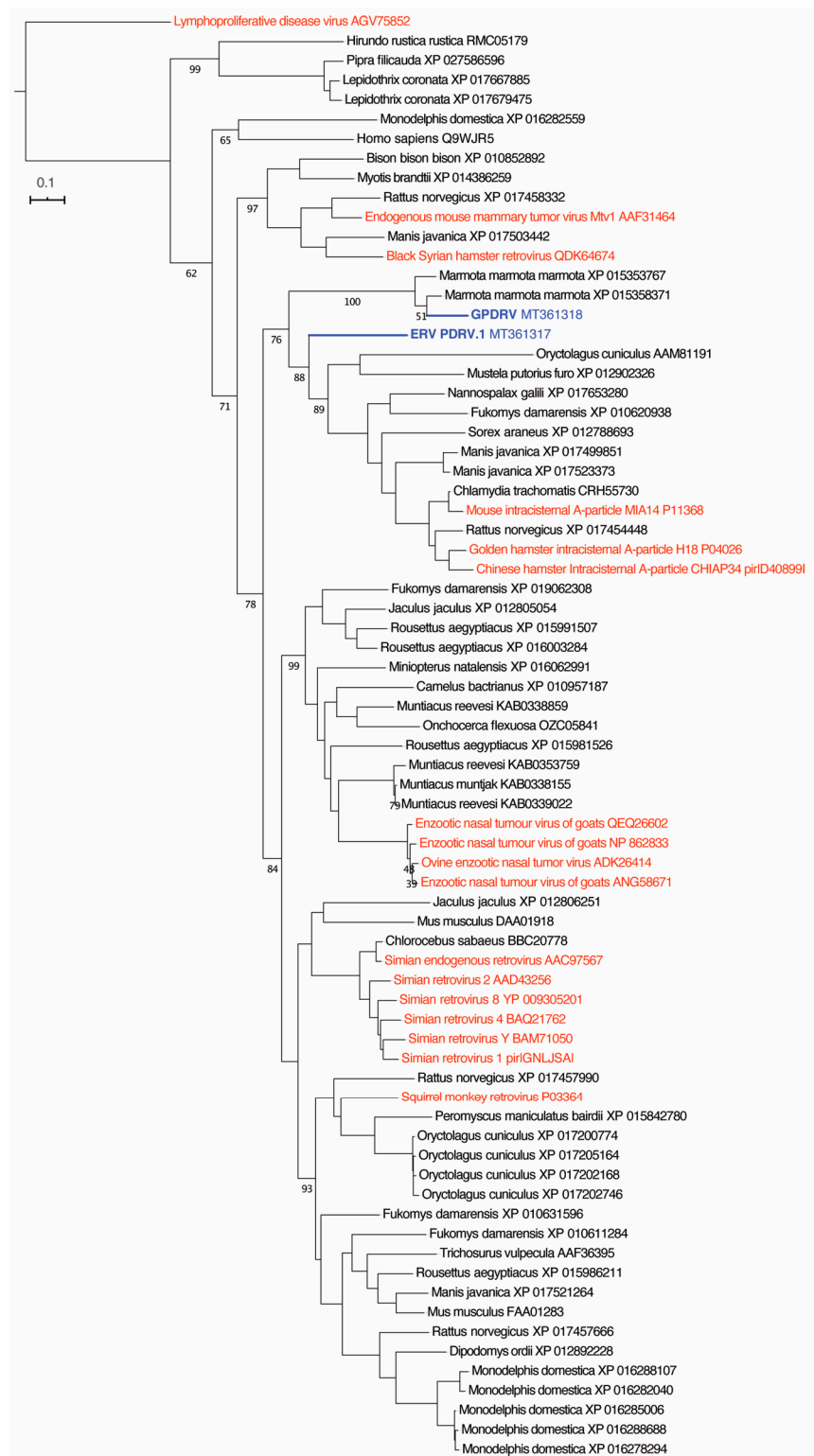
To determine the relationship between GPDRV and members of established retrovirus subfamilies and genera the entire Pol amino acid sequence was used to infer phylogenies. We took a two-fold approach to identify closely related sequences. First, we created a tree using all of the Pol sequences in the NCBI RefSeq protein database that were annotated as belonging to viruses in the family *Retroviridae* (Figure 4). In this tree, GPDRV and ERV- PDRV.1 clustered within betaretrovirus Pol sequences (genus *Betaretrovirus*) (Figure 4). Second, we used BLASTP to identify the most closely related protein sequences in the NCBI protein database: those producing alignments with E-values less than  $10^{-40}$  (Figure 5). The most closely related sequences were a mixture of retroviral and endogenous retroviral-like sequences from mammalian genome assemblies. The GPDRV Pol sequences clustered with sequences present in the alpine marmot (*Marmota marmota*) genome assembly [30]. In fact, GPDRV Pol was more closely related to these marmot sequences than to ERV- PDRV.1 Pol (Figure 5).

### 3.4.5. Env

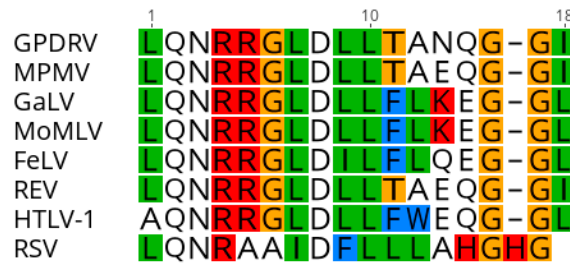
The envelope protein is likely translated from a spliced transcript that utilizes a splice acceptor site located at nucleotide 5209. The *env* open reading frame (nt 5275–7092) is predicted to encode a 605 amino acid, 64.9 kDa protein. A hydrophobic region located from nt 5275 to 5388 would serve as the signal peptide. Proteolytic cleavage at the furin cleavage consensus recognition site (RHRR) in Env would generate a 410 aa, 43.7 kDa surface protein (SU) and 195 aa, 21.2 kDa transmembrane protein (TM). The TM subunit contains two heptad repeats (HR1 and HR2) that form a coiled coil structure. Located between HR1 and HR2 resides a conserved immunosuppressive domain (ISD) (Figure 6) followed by a cysteine-rich region (CX<sub>6</sub>CC), which is predicted to form a covalent disulfide bond with SU. The GPDRV ISD is 94% identical to that found in Mason-Pfizer monkey virus [46]. A 21 aa hydrophobic region within TM (nt 6850–6945) likely serves as the transmembrane anchor with a 43 amino acid cytoplasmic region. There are ten predicted *n*-linked glycosylation sites, nine in SU and one in TM. Phylogenetic analysis of the TM subunit demonstrates GPDRV TM is found in the branch of retroviruses that have an ISD and covalent TM (Figure 7). The group that includes GPDRV TM includes sequences from betaretroviruses and gammaretroviruses, but due to the relatively short length of the TM domain, branches of the tree generally had low support values.



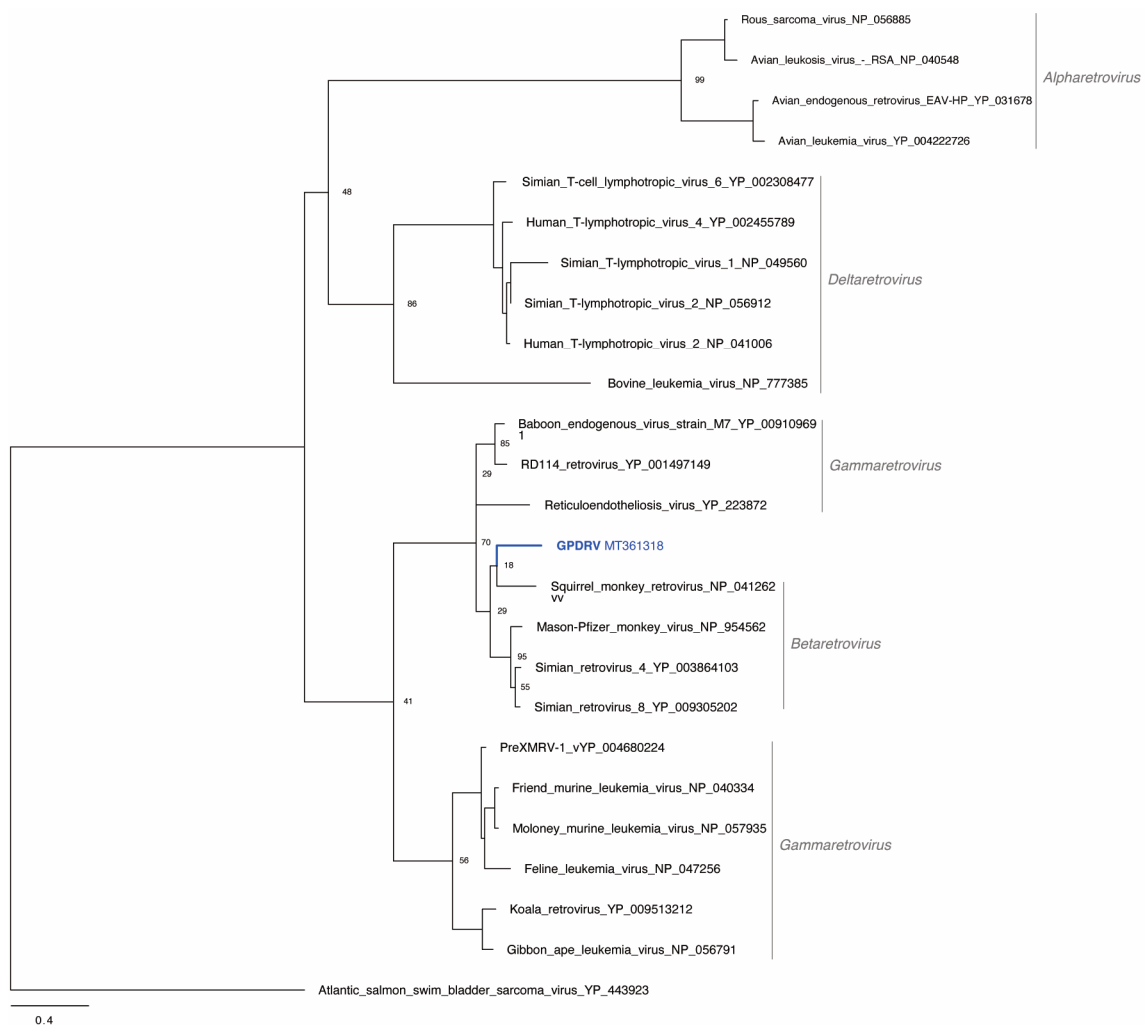
**Figure 4.** GPDRV is a betaretrovirus. All Pol sequences in the NCBI RefSeq protein database annotated as belonging to the *Retroviridae* family longer than 600 amino acids were used to infer a maximum-likelihood tree. The Felsenstein bootstrap (FBP) support values of select branches are indicated. Retrovirus genera and subfamilies (except for *Spumaretrovirinae*) are indicated. The tree is unrooted and was arbitrary midpoint rooted.



**Figure 5.** GPDRV is most closely related to retroviral-like sequences in the marmot genome. A tree was made from the sequences in the NCBI protein database most closely related to GPDRV Pol as determined by a BLASTP search. The Felsenstein bootstrap (FBP) support values of select branches are indicated. The GPDRV and GPD ERV Pol sequences are colored blue. Sequences annotated as belonging to the *Retroviridae* family are colored red. The rest of the sequences are annotated as belonging to the indicated mammalian species and are present in the corresponding genome assemblies. The tree was rooted using lymphoproliferative disease virus Pol, which was included as an outgroup.



**Figure 6.** Alignment of the immunosuppressive domain from retroviral transmembrane protein sequences. Amino acid alignment was constructed using sequences GPDRV- Gunnison’s prairie dog retrovirus (this manuscript), MPMV- Mason-Pfizer monkey virus (NC\_001550), GaLV- Gibbon ape leukemia virus (NC\_001885), MoMLV-Moloney murine leukemia virus (NC\_001501), FeLV- feline leukemia virus (NC\_001940), REV-reticuloendotheliosis virus NC\_006934), HTLV-1- human T cell leukemia virus (NC\_001436), and RSV- Rous sarcoma virus (NC\_001407) with Geneious Prime 2019.2.3.



**Figure 7.** GPDRV Env transmembrane protein (TM) clusters with gamma and betaretrovirus sequences. All ENV TM sequences in the NCBI RefSeq protein database annotated as belonging to the *Retroviridae* family were used to infer a maximum-likelihood tree. The Felsenstein bootstrap (FBP) support values of select branches are indicated. Retrovirus genera are indicated. The tree is unrooted and was arbitrary midpoint rooted.

### 3.5. Integration Site Analysis

Identification of integration sites in tumor DNA resulted in four unique sites; two sites were identified in the 14-1342 tumor and two in the 15-1406 tumor. The Gunnison's prairie dog genome sequence recently became available [47], which enabled identification of the genome location of the four integration sites. The GPDRV sequence in the 14-1342 tumor was found to be integrated at positions 309,469 (+ orientation) and 81,902 (– orientation) and the GPDRV sequence in the 15-1406 tumor was integrated at positions 1,301,104 (+ orientation) and 629,910 (– orientation).

### 3.6. Screening for GPDRV with Virus-Specific Primers

Following the identification of two retroviral sequences with the metagenomics approach we developed virus specific primers to screen additional samples. We selected 42 prairie dog spleen samples to screen for the presence of ERV-PDRV.1-*Cynomys ludovicianus* using virus-specific primers (494F/591R) that amplify a region spanning *pro-pol*. This sequence was detected in DNA from all ( $n = 42$ ) prairie dog spleen samples tested, strongly supporting this as an endogenous retroviral sequence. These samples represented Gunnison's ( $n = 25$ ), black-tailed ( $n = 14$ ), and white-tailed ( $n = 2$ ) prairie dogs.

Alignment of the endogenous and exogenous sequences was used to identify primers that specifically amplify only *gag* from the exogenous GPDRV. Of the 42 prairie dog spleen samples, GPDRV was only detected in the spleen (and tumor) from 14-1342 and 15-1406 and in the spleen of a Gunnison's prairie dog (14-1344) without a tumor. Animal 14-1344 was from the same colony as 14-1342.

Spleen tissues from all 125 prairie dogs in this study were screened for GPDRV using *env*-specific primers 5729F/5840R (Table 1). GPDRV was detected only in Gunnison's prairie dogs, including detection in thymus from 11-1310 and spleen and thymus from 14-1342 and 15-1406. We also detected GPDRV in spleen from four additional Gunnison's prairie dogs that did not have tumors observed at necropsy. Causes of death in these four prairie dogs included: plague (16-633 and 16-778), tularemia (16-675), and suspected stress-related capture mortality (14-1344). Detection of GPDRV RNA is therefore significantly associated with presence of thymic lymphomas (Fisher's exact test;  $p = 0.002$ ).

Using primers located in U3 and *env* we were able to amplify a 7.14 kb fragment from the spleen of animal 14-1344. Sequencing of this product confirmed the presence of GPDRV with 99.7% and 99.1% nucleotide identity with the assembled genomes from 14-1342 and 15-1406, respectively.

## 4. Discussion

In any population, clusters of tumor cases can suggest an underlying or predisposing factor [48–50], and three cases of thymic lymphoma in Gunnison's prairie dogs in Colorado, USA warranted further investigation. We specifically investigated a possible retroviral etiology due to similar retrovirus-associated lymphoid tumors in other species [11,14,15,51,52]. Using PCR and next generation sequencing, we identified and were able to assemble two unique retroviral sequences from lymphoid tissue of prairie dogs.

One of the sequences was identified in DNA from all samples screened, which included members from each of three species of prairie dogs included in the study. This widespread occurrence is consistent with the expected distribution of an endogenous viral sequence [45,53,54]. This sequence lacked an apparent *env*-coding region, suggesting a possible mutation/deletion typical of endogenous viruses [53,54]. We suspect that this consensus sequence likely represents endogenous retroviral sequence(s) of prairie dogs, and we propose the name ERV-PDRV.1-*Cynomys ludovicianus* [55]. The presence of ERV-PDRV.1-*Cynomys ludovicianus* in all three species of prairie dogs suggests that endogenization occurred prior to evolutionary divergence of these species [45,53,56]. Screening of other prairie dog species and other *Sciuridae* species may provide further insights as to when this viral sequence was acquired. Gifford et al. [45] screened for the presence of class II endogenous retroviral sequences using conserved PR and RT primers and identified a sequence from a black-tailed prairie dog

that is 62% identical to the ERV-PDRV.1-*Cynomys ludovicianus* and 59% identical to GPDRV nucleotide sequences identified in this study.

A second retroviral sequence identified in thymic tissue and spleen of two prairie dogs with thymic lymphoma included *gag-pro-pol-env* coding regions. Using *env*-specific primers this sequence was identified in thymic tissue and spleen from all (3/3) of the prairie dogs with thymic lymphoma, and from splenic tissue of only 3.3% (4/122) of prairie dogs that did not have tumors identified grossly. All of the *env*-positive animals were from prairie dog colonies located in the Gunnison Basin in Colorado, USA. Gunnison's prairie dogs located outside the Gunnison Basin were all negative for GPDRV. This consistent association with tumors and infrequent occurrence in the overall population of animals examined in the study is an expected pattern for an infectious, exogenous oncogenic virus. Isolation of virus in culture was not pursued in this study. The finding of multiple integration sites supports classification as an exogenous infectious virus. We suggest that this sequence represents the first exogenous retroviral sequence identified in prairie dogs and propose the name Gunnison's prairie dog retrovirus (GPDRV).

As we were finalizing this paper, a Gunnison's prairie dog genome assembly was published [47]. This assembly contained multiple contigs with sequences similar to both GPRDV and to ERV-PDRV.1-*Cynomys ludovicianus*. A BLASTN search of the assembly with ERV-PDRV.1-*Cynomys ludovicianus* yielded 3156 alignments with *E*-values  $<1e-10$ . Many of these are nearly identical to the ERV-PDRV.1 sequence over more or less its entire length: 652 of the alignments cover  $>80\%$  of the ERV sequence with  $>90\%$  identity. The highest scoring alignment was 95.9% identical over 100% of the ERV-PDRV.1 sequences, which represents a consensus sequence assembled from our metagenomic datasets. A similar BLAST search with GPDRV produced 759 alignments with *E*-values  $<10^{-10}$ . For these, the highest scoring alignment was on 84.9% identical to GPDRV, over 99% of the GPDRV sequence. We conclude that the Gunnison's prairie dog genome contains a large number of sequences related to both GPDRV and to ERV-PDRV.1 that may represent endogenized retrovirus sequences and possibly proviruses from other exogenous retroviruses.

The genetic organization of GPDRV is typical of a betaretrovirus. The protease and polymerase proteins are expected to be expressed as Gag-Pro and Gag-Pro-Pol polypeptides by ribosomal frameshifting. The protein encoded by *pro* harbors dUTPase and active protease domains. Phylogenetic analysis based on Pol amino acid sequences supported the classification of GPDRV as a betaretrovirus.

The retrovirus envelope protein is cleaved by cellular furin to generate SU and TM subunits. The SU and TM subunits remain associated after cleavage either through noncovalent interactions or formation of a covalent bond, which is determined by the cysteine motif located between the heptad repeats within TM [57,58]. The betaretroviruses, MMTV, JSRV and ENTV, contain a CX<sub>7</sub>C motif, which forms noncovalent interactions [59]. The TM subunit found in alpha-, gamma- and delta- retroviruses contain a cysteine motif, CX<sub>6</sub>CC that forms a covalent bond with a cysteine in the SU subunit [57,60–62]. The TM subunit of alpha-, gamma-, delta-retroviruses and the D-type betaretrovirus, Mason-Pfizer monkey virus (MPMV), also contain an immunosuppressive domain [46,63,64]. The envelope gene of MPMV was derived by a recombination event, which resulted in the acquisition of an envelope gene from a gammaretrovirus [46,65]. The GPDRV TM harbors a CX<sub>6</sub>CC and an immunosuppressive domain similar to that of MPMV suggesting GPDRV is a D-type betaretrovirus that may have undergone a similar recombination event.

The significance of GPDRV to prairie dog populations is uncertain. For all three cases with thymic lymphoma, the tumor was determined to be the cause of death or related to the cause of death by compromising cardiovascular function under anesthesia. However, overall occurrence of thymic lymphoma was low (3/125) and not considered to be a significant source of mortality at the population level. Of the four prairie dogs which tested positive for GPDRV but were unaffected by thymic lymphoma, three died from bacterial infections. Based on the immunosuppressive effects of other retroviruses, we considered possible population effects due to increased susceptibility to bacterial infections. However, prairie dogs positive for GPDRV were not over-represented among animals that died from bacteremia. Although three of the seven (43%) prairie dogs with GPDRV died

from bacterial infections, 51 of the 118 (43%) prairie dogs without GPDRV also died from bacterial infections. One prairie dog that was positive for GPRDV but unaffected by thymic lymphoma died from capture-related factors (heat stress) with no signs of immune suppression. This rare complication of trapping is unfortunate but GPDRV was not suspected to be associated with the cause of death in this animal. No other lesions observed were suggestive of immune suppression. The small proportion of animals testing positive for GPLDRV prevented further analysis of possible effects of the virus.

GPDRV is statistically associated with thymic lymphoma in Gunnison's prairie dogs. Future investigations could include isolation of virus, experimental infection studies, and corroboration of the association with larger sample sets. Further analysis of integration sites may yield insight into potential mechanisms of oncogenesis. Surveillance areas were limited to the state of Colorado, with access to only three prairie dog species. GPDRV sequences were only found in Gunnison's prairie dogs. This could suggest either species specificity or lack of exposure in other species. Continued surveillance of other prairie dog species may help understand the species host range of the virus.

**Author Contributions:** Conceptualization, S.L.Q., M.D.S., and K.A.F.; methodology, S.L.Q., M.D.S., and K.A.F.; software, M.D.S.; validation, S.L.Q., M.D.S., and K.A.F.; formal analysis M.D.S., M.D.B., S.L.Q., and K.A.F.; investigation, M.D.B., K.G., C.D.B., M.L.K., and K.A.F.; resources, D.W.T., K.A.F., and M.D.S.; data curation, M.L.K. and M.D.S.; writing—original draft preparation, K.A.F., M.D.B., S.L.Q., and M.D.S.; writing—review and editing, M.D.B., S.L.Q., M.D.S., M.L.K., D.W.T., and K.A.F.; visualization, M.D.B., S.L.Q., K.A.F., M.D.S., and M.L.K.; supervision, S.L.Q., M.D.S., and K.A.F.; project administration, S.L.Q., M.D.S., and K.A.F.; funding acquisition, S.L.Q., M.D.S., and D.W.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by funds from the College of Veterinary Medicine and Biomedical Sciences (S.L.Q.). Informatics was supported by NIH/NCATS Colorado CTSA Grant Number UL1 TR002535 (M.D.S.). Colorado Parks and Wildlife and Colorado's Species Conservation Trust Fund supported fieldwork (D.W.T.).

**Acknowledgments:** Thanks also to Michael Miller for consultation and review of previous drafts of this manuscript, and to the many field technicians who provided assistance and technical support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hardwicke, K. Prairie Dogs, Plants, and Pollinators: Tritrophic Interactions Affect Plant-Insect Floral Visitor Webs in Shortgrass Steppe. Colorado State University: Fort Collins, CO, USA, 20 July 2006; Available online: <https://search.proquest.com/docview/305356392> (accessed on 7 May 2020).
2. Kotliar, n.B.; Miller, B.J.; Reading, R.P.; Clark, T.W. *The Prairie Dog as a Keystone Species*; Island Press: Washington, DC, USA, 2006; pp. 53–64.
3. Richgels, K.L.; Russell, R.E.; Bron, G.M.; Rocke, T.E. Evaluation of Yersinia pestis Transmission Pathways for Sylvatic Plague in Prairie Dog Populations in the Western U.S. *Ecohealth* **2016**, *13*, 415–427. [[CrossRef](#)] [[PubMed](#)]
4. Ecke, D.H.; Johnson, C.W. Part I. Plaque in Colorado. In *Plague in Colorado, Public Health Monograph No. 6*; US Government Printing Office: Washington, DC, USA, 1952; pp. 1–37.
5. Thas, I.; Garner, M.M. A retrospective study of tumours in black-tailed prairie dogs (*Cynomys ludovicianus*) submitted to a zoological pathology service. *J. Comp. Pathol.* **2012**, *147*, 368–375. [[CrossRef](#)] [[PubMed](#)]
6. Wright, T.L.; Eshar, D.; Carpenter, J.W.; Lin, D.; Padmanabhan, A.; Peddireddi, L.; Cino, G. Suspected Hepadnavirus Association with a Hepatocellular Carcinoma in a Black-Tailed Prairie Dog (*Cynomys ludovicianus*). *J. Comp. Pathol.* **2017**, *157*, 284–290. [[CrossRef](#)] [[PubMed](#)]
7. Matsumoto, I.; Chambers, J.K.; Miwa, Y.; Nakayama, H.; Uchida, K. Anaplastic large T-cell lymphoma in three black-tailed prairie dogs (*Cynomys ludovicianus*). *J. Vet. Med. Sci.* **2017**, *79*, 979–983. [[CrossRef](#)] [[PubMed](#)]
8. Miwa, Y.; Matsunaga, S.; Nakayama, H.; Kurosawa, A.; Ogawa, H.; Sasaki, N. Spontaneous lymphoma in a prairie dog (*Cynomys ludovicianus*). *J. Am. Anim. Hosp. Assoc.* **2006**, *42*, 151–153. [[CrossRef](#)]
9. Ellermann, V.; Bang, O. Experimentelle Leukämie bei Hühnern. *Zentralbl. Bakteriologie. Parasitenkd. Infektionskr. Hyg. Abt. Hygiene* **1908**, *46*, 595–609. [[CrossRef](#)]
10. Hardy, W.D.; Old, L.J.; Hess, P.W.; Essex, M.; Cotter, S. Horizontal transmission of feline leukaemia virus. *Nature* **1973**, *244*, 266–269. [[CrossRef](#)]



11. Jarrett, W.F.; Martin, W.B.; Crichton, G.W.; Dalton, R.G.; Stewart, M.F. Leukemia in the cat. Transmission experiments with leukemia (lymphosarcoma). *Nature* **1964**, *202*, 566. [[CrossRef](#)]
12. Mammerickx, M.; Portetelle, D.; de Clercq, K.; Burny, A. Experimental transmission of enzootic bovine leukosis to cattle, sheep and goats: Infectious doses of blood and incubation period of the disease. *Leuk. Res.* **1987**, *11*, 353–358. [[CrossRef](#)]
13. Poiesz, B.J.; Ruscetti, F.W.; Gazdar, A.F.; Bunn, P.A.; Minna, J.D.; Gallo, R.C. Detection and isolation of type-C retrovirus particles from fresh and cultured lymphocytes of patients with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 7415–7419. [[CrossRef](#)]
14. Rosenberg, N.; Jolicoeur, P. *Retroviral Pathogenesis*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 1997.
15. Xu, W.; Stadler, C.K.; Gorman, K.; Jensen, n.; Kim, D.; Zheng, H.; Tang, S.; Switzer, W.M.; Pye, G.W.; Eiden, M.V.; et al. An exogenous retrovirus isolated from koalas with malignant neoplasias in a US zoo. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11547–11552. [[CrossRef](#)]
16. Beatty, J.A.; Lawrence, C.E.; Callanan, J.J.; Grant, C.K.; Gault, E.A.; Neil, J.C.; Jarrett, O. Feline immunodeficiency virus (FIV)-associated lymphoma: A potential role for immune dysfunction in tumorigenesis. *Vet. Immunol. Immunopathol.* **1998**, *65*, 309–322. [[CrossRef](#)]
17. Hartmann, K. Clinical aspects of feline retroviruses: A review. *Viruses* **2012**, *4*, 2684–2710. [[CrossRef](#)]
18. Hoover, E.A.; Mullins, J.I.; Quackenbush, S.L.; Gasper, P.W. Experimental transmission and pathogenesis of immunodeficiency syndrome in cats. *Blood* **1987**, *70*, 1880–1892. [[CrossRef](#)]
19. Magden, E.; Miller, C.; MacMillan, M.; Bielefeldt-Ohmann, H.; Avery, A.; Quackenbush, S.L.; Vandewoude, S. Acute virulent infection with feline immunodeficiency virus (FIV) results in lymphomagenesis via an indirect mechanism. *Virology* **2013**, *436*, 284–294. [[CrossRef](#)]
20. Orosz, C.G.; Zinn, n.E.; Olsen, R.G.; Mathes, L.E.; Zinn, N.E.; Olsen, R.G.; Mathes, L.E. Retrovirus-mediated immunosuppression. I. FeLV-UV and specific FeLV proteins alter T lymphocyte behavior by inducing hyporesponsiveness to lymphokines. *J. Immunol.* **1985**, *134*, 3396–3403.
21. Tripp, D.W.; Rocke, T.E.; Runge, J.P.; Abbott, R.C.; Miller, M.W. Burrow Dusting or Oral Vaccination Prevents Plague-Associated Prairie Dog Colony Collapse. *Ecohealth* **2017**, *14*, 451–462. [[CrossRef](#)]
22. Begier, E.M.; Asiki, G.; Anywaine, Z.; Yockey, B.; Schriefer, M.E.; Aleti, P.; Ogden-Odoi, A.; Staples, J.E.; Sexton, C.; Bearden, S.W.; et al. Pneumonic plague cluster, Uganda, 2004. *Emerg. Infect Dis.* **2006**, *12*, 460–467. [[CrossRef](#)]
23. Griffin, K.A.; Martin, D.J.; Rosen, L.E.; Sirochman, M.A.; Walsh, D.P.; Wolfe, L.L.; Miller, M.W. Detection of *Yersinia pestis* DNA in prairie dog-associated fleas by polymerase chain reaction assay of purified DNA. *Journal of wildlife diseases* **2010**, *46*, 636–643. [[CrossRef](#)]
24. Junhui, Z.; Ruifu, Y.; Jianchun, L.; Songle, Z.; Meiling, C.; Fengxiang, C.; Hong, C. Detection of *Francisella tularensis* by the polymerase chain reaction. *J. Med. Microbiol.* **1996**, *45*, 477–482. [[CrossRef](#)]
25. Donehower, L.A.; Bohannon, R.C.; Ford, R.J.; Gibbs, R.A. The use of primers from highly conserved pol regions to identify uncharacterized retroviruses by the polymerase chain reaction. *J. Virol. Methods* **1990**, *28*, 33–46. [[CrossRef](#)]
26. Caspritz, G.; Hadden, J. The immunopharmacology of immunotoxicology, and immunorestitution. *Tox. Pathol.* **1989**, *15*, 320–332. [[CrossRef](#)]
27. Cross, S.T.; Kapuscinski, M.L.; Perino, J.; Maertens, B.L.; Weger-Lucarelli, J.; Ebel, G.D.; Stenglein, M.D. Co-Infection Patterns in Individual *Ixodes scapularis* Ticks Reveal Associations between Viral, Eukaryotic and Bacterial Microorganisms. *Viruses* **2018**, *10*, 388. [[CrossRef](#)]
28. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12. [[CrossRef](#)]
29. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)]
30. Gossmann, T.I.; Shanmugasundram, A.; Borno, S.; Duvaux, L.; Lemaire, C.; Kuhl, H.; Klages, S.; Roberts, L.D.; Schade, S.; Gostner, J.M.; et al. Ice-Age Climate Adaptations Trap the Alpine Marmot in a State of Low Genetic Diversity. *Curr. Biol.* **2019**, *29*, 1712–1720.e7. [[CrossRef](#)]
31. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]

32. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)]
33. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **2015**, *12*, 59–60. [[CrossRef](#)]
34. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421. [[CrossRef](#)]
35. RC Team. *R: A language and environment for statistical computing*; R foundation for Statistical Computing: Vienna, Austria, 2017.
36. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
37. Capella-Gutierrez, S.; Silla-Martinez, J.M.; Gabaldon, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [[CrossRef](#)]
38. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
39. Darriba, D.; Posada, D.; Kozlov, A.M.; Stamatakis, A.; Morel, B.; Flouri, T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.* **2020**, *37*, 291–294. [[CrossRef](#)]
40. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
41. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]
42. Cousens, C.; Minguignon, E.; Dalziel, R.G.; Ortin, A.; Garcia, M.; Park, J.; Gonzalez, L.; Sharp, J.M.; de las Heras, M. Complete sequence of enzootic nasal tumor virus, a retrovirus associated with transmissible intranasal tumors of sheep. *J. Virol.* **1999**, *73*, 3986–3993. [[CrossRef](#)]
43. York, D.F.; Vigne, R.; Verwoerd, D.W.; Querat, G. Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats. *J. Virol.* **1992**, *66*, 4930–4939. [[CrossRef](#)]
44. Aravind, L.; Koonin, E.V. G-patch: A new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends Biochem. Sci.* **1999**, *24*, 342–344. [[CrossRef](#)]
45. Gifford, R.; Kabat, P.; Martin, J.; Lynch, C.; Tristem, M. Evolution and distribution of class II-related endogenous retroviruses. *J. Virol.* **2005**, *79*, 6478–6486. [[CrossRef](#)]
46. Sonigo, P.; Barker, C.; Hunter, E.; Wain-Hobson, S. Nucleotide sequence of Mason-Pfizer monkey virus: An immunosuppressive D-type retrovirus. *Cell* **1986**, *45*, 375–385. [[CrossRef](#)]
47. Tsuchiya, M.T.n.; Dikow, R.B.; Cassin-Sackett, L.N.; Dikow, R.B.; Cassin-Sackett, L. First genome sequence of the Gunnison’s prairie dog (*Cynomys gunnisoni*), a keystone species and player in the transmission of sylvatic plague. *Genome Biol. Evol.* **2020**. [[CrossRef](#)] [[PubMed](#)]
48. Chang, Y.; Moore, P.S.; Weiss, R.A. Human oncogenic viruses: Nature and discovery. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2017**, *372*. [[CrossRef](#)] [[PubMed](#)]
49. Hardy, W.D., Jr.; McClelland, A.J.; MacEwen, E.G.; Hess, P.W.; Hayes, A.A.; Zuckerman, E.E. The epidemiology of the feline leukemia virus (FeLV). *Cancer* **1977**, *39*, 1850–1855. [[CrossRef](#)]
50. Parkin, D.M. The global health burden of infection-associated cancers in the year 2002. *Int. J. Cancer* **2006**, *118*, 3030–3044. [[CrossRef](#)]
51. Hoover, E.A.; Rojko, J.L.; Olsen, R.G. Pathogenesis of feline leukemia virus infection. In *Feline Leukemia*; Olsen, R.G., Ed.; CRC Press: Boca Raton, FL, USA, 1980; pp. 32–51.
52. Kozak, C.A.; Ruscetti, S. Retroviruses in Rodents. In *The Retroviridae*; Levy, J.A., Ed.; Plenum Press: New York, NY, USA, 1992; Volume 1, pp. 405–481.
53. Boeke, J.D.; Stoye, J.P. Retrotransposons, Endogenous Retroviruses, and the Evolution of Retroelements. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, H.E., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor, NY, USA, 1997.
54. Mager, D.L.; Stoye, J.P. Mammalian Endogenous Retroviruses. *Microbiol. Spectr* **2015**, *3*, MDNA3-0009-2014. [[CrossRef](#)]
55. Gifford, R.J.; Blomberg, J.; Coffin, J.M.; Fan, H.; Heidmann, T.; Mayer, J.; Stoye, J.; Tristem, M.; Johnson, W.E. Nomenclature for endogenous retrovirus (ERV) loci. *Retrovirology* **2018**, *15*, 59. [[CrossRef](#)]

56. Arnaud, F.; Varela, M.; Spencer, T.E.; Palmarini, M. Coevolution of endogenous betaretroviruses of sheep and their host. *Cell Mol. Life Sci.* **2008**, *65*, 3422–3432. [[CrossRef](#)]
57. Opstelten, D.J.; Wallin, M.; Garoff, H. Moloney murine leukemia virus envelope protein subunits, gp70 and Pr15E, form a stable disulfide-linked complex. *J. Virol.* **1998**, *72*, 6537–6545. [[CrossRef](#)]
58. Wallin, M.; Ekstrom, M.; Garoff, H. Isomerization of the intersubunit disulphide-bond in Env controls retrovirus fusion. *EMBO J.* **2004**, *23*, 54–65. [[CrossRef](#)]
59. Henzy, J.E.; Coffin, J.M. Betaretroviral envelope subunits are noncovalently associated and restricted to the mammalian class. *J. Virol.* **2013**, *87*, 1937–1946. [[CrossRef](#)] [[PubMed](#)]
60. Johnston, E.R.; Radke, K. The SU and TM envelope protein subunits of bovine leukemia virus are linked by disulfide bonds, both in cells and in virions. *J. Virol.* **2000**, *74*, 2930–2935. [[CrossRef](#)] [[PubMed](#)]
61. Leamson, R.n.; Halpern, M.S.N.; Halpern, M.S. Subunit structure of the glycoprotein complex of avian tumor virus. *J. Virol.* **1976**, *18*, 956–968. [[CrossRef](#)] [[PubMed](#)]
62. Li, K.; Zhang, S.; Kronqvist, M.; Wallin, M.; Ekstrom, M.; Derse, D.; Garoff, H. Intersubunit disulfide isomerization controls membrane fusion of human T-cell leukemia virus Env. *J. Virol.* **2008**, *82*, 7135–7143. [[CrossRef](#)] [[PubMed](#)]
63. Benit, L.; Dessen, P.; Heidmann, T. Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. *J. Virol.* **2001**, *75*, 11709–11719. [[CrossRef](#)]
64. Cianciolo, G.J.; Copeland, T.D.; Oroszlan, S.; Snyderman, R. Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. *Science* **1985**, *230*, 453–455. [[CrossRef](#)]
65. Barker, C.S.; Pickel, J.; Tainsky, M.; Hunter, E. Molecular cloning of the Mason-Pfizer monkey virus genome: Biological characterization of genome length clones and molecular comparisons to other retroviruses. *Virology* **1986**, *153*, 201–214. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).