A First NGS Investigation Suggests No Association Between Viruses and Canine Cancers

Diana Giannuzzi[1] and Luca Aresu[2*]

[1] Department of Comparative Biomedicine and Food Science, University of Padua, Legnaro, Italy, [2] Department of Veterinary Science, University of Turin, Grugliasco, Italy

Approximately 10–15% of worldwide human cancers are attributable to viral infection. When operating as carcinogenic elements, viruses may act with various mechanisms, but the most important is represented by viral integration into the host genome, causing chromosome instability, genomic mutations, and aberrations. In canine species, few reports have described an association between viral integration and canine cancers, but more comprehensive studies are needed. The advancement of next-generation sequencing and the cost reduction have resulted in a progressive increasing of sequencing data in veterinary oncology offering an opportunity to study virome in canine cancers. In this study, we have performed viral detection and integration analyses using VirusFinder2 software tool on available whole-genome and whole-exome sequencing data of different canine cancers. Several viral sequences were detected in lymphomas, hemangiosarcomas, melanomas, and osteosarcomas, but no reliable integration sites were identified. Even if with some limitations such as the depth and type of sequencing, a restricted number of available nonhuman genomes software, and a limited knowledge on endogenous retroviruses in the canine genome, results are compelling. However, further experiments are needed, and similarly to feline species, dedicated analysis tools for the identification of viral integration sites in canine cancers are required.

Keywords: animal model, canine tumors, NGS, oncogenesis, viral integration

INTRODUCTION

An estimated 10–15% of all human cancers are attributed to viruses, including papillomavirus (PV), hepatitis B virus (HBV), Epstein–Barr virus (EBV), human T-lymphotropic virus, Kaposi sarcoma–associated herpesvirus, and Merkel cell polyomavirus (1–6). Although individual tumor viruses exert their oncogenic effects in different ways, a common feature is represented by the integration of the viral genome into the host, causing gene mutations, chromosome instability, and genomic aberrations (7, 8). In canine species, only few reports have pinpointed an association between viruses and tumors, and canine PV integration was described both in squamous cell carcinoma and transmissible venereal tumor (9–12). Although the presence alone of viral sequences in tumor genome is insufficient to prove an oncogenic role, viral integration may represent the first hallmark of the malignant transformation. Indeed, interplay between viral variants and proto-oncogenes conferring malignancy needs to be proved. This is the case of gammaretrovirus feline leukemia virus (FeLV) that is widely recognized to be oncogenic, predisposing cats to lymphoma, and fibrosarcoma (13). The pathogenic mechanisms of FeLV influencing lymphoma transformation

were previously identified and include common integration sites nearby proto-oncogenes, resulting in insertional mutagenesis (14). Conversely, the association of feline immunodeficiency virus (FIV) in feline lymphoma was reported as a rare event (15, 16), and the pathogenetic role is scarcely known. Furthermore, FIV-related malignant transformation seems to imply different mechanisms from those employed by FeLV, and a causal association has been described only when provirus integration in tumor DNA is monoclonal (17).

The advent of next-generation sequencing (NGS) has caused a tremendous impact both on basic and clinical oncology research. Next-generation sequencing analysis of tumor genetic changes, transcriptomes, and epigenomes is driving biomarker discovery for cancer diagnostic and therapy, but it also opens opportunities for the identification of new associations between tumors and viruses (2, 8, 18). Oncoviral profiles are obtained from cancer genome sequencing data, and several algorithms have been developed to identify viral integration sites (19–23), but only a few are available for nonhuman host genomes (19, 24). In any case, the computational approaches are very similar and based on the identification of one or more reads containing a portion of nucleic acid sequences uniquely mapping to a viral genome integrated within tumoral DNA. As costs become less prohibitive, and methods simpler, veterinary researchers are choosing NGS over microarrays to study canine tumors, and the recent studies published on this topic provide an unprecedented opportunity to study canine virome.

Considering these premises, our study aims at (i) detecting the presence of viruses of arbitrary types in NGS data from several canine tumor histotypes; (ii) identifying virus integration sites as long as viruses merge into the genome and the sequencing captures the insertion sites; and (iii) evaluating the potential pathogenic role of the viral presence/integration.

METHODS

First, we have performed a systematic review to retrieve the current literature where NGS technology was applied to canine tumors, and if publicly available, we downloaded FASTQ files of paired-end whole-exome sequencing (WES) and whole-genome sequencing (WGS) using the SRA toolkit (https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/). The dataset included 107 lymphomas (25), three transitional cell carcinomas (TCCs) of the urinary bladder (26), 20 splenic hemangiosarcomas (27), 11 malignant melanomas (28), 18 acanthomatous ameloblastomas (29), 46 osteosarcomas (30), 22 gliomas (31), 20 mammary tumors (32), and 28 mast cell tumors (33) (**Table S1**). The mean sequencing coverage of each dataset is reported in **Table S2**. The viral detection, integration, and mutation identification were performed using VirusFinder2 software tool implemented by virus integration site detection through reference sequence customization (VERSE) (19, 24). Briefly, VirusFinder2 integrates a mixed strategy of NGS reads alignment. First, raw reads are aligned to the host reference genome, and then the unmapped are aligned to a hybrid reference genome for virus detection and integration (34). In this study, all the sequences were aligned to

the dog genome (CanFam3.1, Sept. 2011, Ensembl release 96). The virus genomes were obtained from RINS package (35) and National Center for Biotechnology Information (NCBI) virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/). In details, RINS package includes 32,102 viral genomes derived from all the classes in GenBank (36) and International Committee on Taxonomy of Viruses retrieved through the NCBI. We included all partial and complete sequences ($n = 16,394$) of known viruses having *Canis lupus familiaris* as a putative host from the NCBI virus database. The configuration file was set with default parameters, except for "sensitivity_level = 2" and "min_contig length = 200," which were modified to reach a higher sensitivity during virus detection. The detailed list of third-party tools used by VirusFinder2 and an example of configuration file are provided in **Supplementary Methods**.

RESULTS

After the preprocessing step, the pipeline of VirusFinder2 follows a two-step procedure: (1) virus detection and (2) virus integration site detection. The virus detection did not retrieve viral sequences in TCCs and ameloblastomas. Conversely, several putative viral sequences were aligned against host unmapped sequences with high identity (>85%) in mammary tumors ($n = 4$), gliomas ($n = 1$), mast cell tumors ($n = 7$), lymphomas ($n = 8$), hemangiosarcomas ($n = 6$), melanomas ($n = 11$), and osteosarcomas ($n = 27$). By definition, the identity indicates similarity between host unmapped and viral sequences. Further, the data were manually checked for common bacteria and bacteriophage sequences and when identified were excluded as result of inaccurate annotation (37). Also, NGS computational artifacts and common viral contaminants including *Choristoneura occidentalis* granulovirus and *Paramecium bulsaria chlorella* virus were discarded (38, 39). The host genome unmapped reads fallen on contigs lower than 1,000 were excluded.

After this filtering process, all the candidate viruses in gliomas, hemangiosarcomas, and mammary tumors resulted false positives. Instead, 1 lymphoma, 7 mast cell tumors, and 12 osteosarcomas retained several candidate viruses. The top-ranking virus candidate was investigated for each tumor based on percentage of identity, length of the contigs, and number of reads fallen on contigs (**Table S3**). Then, the viral integration was tested, but only three osteosarcomas showed a positive result. Viruses identified were Meleagrid herpesvirus (HV) 1, Gallid HV 3, and Human HV 7 (**Table 1**), but a low confidence was assessed.

Because several herpesviruses were retrieved among the top-ranking viruses, we manually checked the correspondences. All the sequences aligned to unspecific homopolymer-rich sequences, confirming previously published data reporting that herpesviruses harbor telomeric repeats almost identical to mammalian telomeres (40–42). Also, a retrovirus was included in the list, but the sequence resulted in an long terminal repeat (LTR)-rich sequence, likely corresponding to proviral DNA or endogenous retroviruses (ERVs).

TABLE 1 | List of virus integration sites in all the cancer samples.

Sample (1)	Chr (1)	Position (1)	Virus (2)	Chr (2)	Position (2)	No. support reads (pair + softclip)	Confidence
SRR8691727	chr12	72,497,897	AF282130.1__Meleagrid_herpesvirus_1	Viral genome	117,487–117,676	31 + 9	Low
SRR8691785	chr12	72,497,897	NC_002577.1__Gallid_herpesvirus_3	Viral genome	133,794–134,005	73 + 18	Low
SRR8691786	chr12	72,497,897	NC_001716.2__Human_herpesvirus_7	Viral genome	143,544–143,960	27 + 11	Low

Chr, chromosome; 1, attributes related to the cancer sample; 2, attributes related to the virus; pair + softclip, the word pair means the number of paired-end reads whose one end mapped to the host genome and another aligned to the integrated virus sequence. The word softclip means the number of reads that actually harbor the integration breakpoint within themselves. Softclip is the total soft-clipped reads that map to the host reference genome plus the total soft-clipped reads that aligned to the virus sequence; confidence: if the number of soft-clipped reads is sufficient to support an accurate characterization of the locus.

Last, to get insight into the putative viral sequences of lymphoma, mast cell tumor, and osteosarcoma samples, the GenBank nonredundant database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/) was interrogated. Correspondences were obtained only in osteosarcoma, mainly matching with *Macrostomum lignano* and *Branchiostoma floridae*, although identity was overall low (<55%).

DISCUSSION

In recent years, a significant amount of NGS data has been produced in dogs, but no studies were conducted to investigate virome in cancers. Also, several algorithms have been developed for NGS investigation of virus–host interactions in human species. In this study, we selected VirusFinder2, which resulted the most performant software by far. This was similarly demonstrated in a previous study comparing methods for the identification of viral integrations in genomes (34). Indeed, VirusFinder2 was tested on several human cancer genomes (19, 24) and showed the highest sensitivity and precision in high-coverage datasets (>60×) compared to other software able to handle multiple-virus integration. In our dataset, the tumors were sequenced with a mean coverage higher than 60×, with the exception of the TCC and melanoma WGS samples (**Table S2**). Further, VirusFinder2 integrates nonhuman genomes representing a significant advantage in veterinary medicine.

Given these premises, in our study, we tested a large WGS and WES canine tumors dataset, but none of the detected virus sequences and viral integration sites were reliable, and this is determined by the low confidence of the integration results. Indeed, all the detected viruses do not infect the canine species, and the candidate sequences were mostly unspecific homopolymer-rich sequences of herpesviruses or LTR-rich sequences of retroviruses. Few sequences also aligned with genomes of marine organisms, suggesting a contamination of the tumor samples.

Interestingly, several sequences were suspected to correspond to ERVs. Endogenous retroviruses show a ubiquitous presence in vertebrate genomes comprising up to 5–8% of the human genome. While the functional effects of ERVs are well-established, and the vast majority of ERV insertions are fixed across individuals, the potential contribution to phenotypic variation is still unknown in dog. Also, degeneration due to the accumulation of mutations can occur during life

(43, 44). In cats infected by FeLV, a recombination between endogenous and exogenous viruses can generate strains with specific pathogenic capabilities (45). The canine genome displays only a low percentage (∼0.15%) of ERVs compared to other mammals (46, 47), and this is mainly related to a low number of repetitive elements in the genome (48). Despite hematological malignancies are frequent in dogs, and retroviral particles have been reported both in lymphomas and cell lines (49, 50), no exogenous retroviruses have been identified in dogs or any other canid so far. The lack of exogenous retroviruses and the low number of ERVs are quite surprising mainly because canine cells are known to enable the replication of noncanid exogenous retroviruses, and the canine genome lacks the functional retroviral restriction factor TRIM5α (51). Recent studies have proposed that the expansion of an endogenous gammaretroviral lineage might originate from an infection in canid ancestors, thereby supporting the potential existence of canine exogenous retroviruses (46, 52). Further studies are needed to confirm these findings.

A partial explanation for the negative results obtained here may be the selected dataset that comprised mainly WES data, implying that sequencing targets covered only exonic regions. Although retroviruses can integrate randomly or in transcription units, introns and exons of active genes (5), the majority of retroviruses, PVs, HBV, and EBV, favor integration into promoters and transcription start sites of active genes (53–55) that are rarely sequenced in WES. Here, the integration sites were retrieved only in WGS samples.

Finally, although this is not the case, the presence alone of viral sequences in tumor genome is insufficient to prove an oncogenic role. In humans, several viruses are widely distributed in the population such as EBV and human HV 6, but not causing malignancies. A more complicated aspect to solve is the discrimination of the time point at which viral infections may have arisen: if infections occur after tumor growth, then they have no contribution to tumorigenesis. Unfortunately, we could not expand this aspect considering the negative results.

In conclusion, we have analyzed WES and WGS data derived from different canine tumor histotypes to detect potential viral integration sites. Even if previous lines of evidence have hypothesized a role of viruses in canine lymphoma (56), we could not confirm this hypothesis. However, similarly to feline species (57), studies dedicated to identify viral integration sites in canine cancers are needed, and transmissible venereal tumor

and squamous carcinoma represent excellent models to further test VirusFinder2.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: PRJNA389294, PRJNA478811, PRJNA417727, PRJNA525883, PRJNA516699, PRJNA339175, PRJNA247493, and PRJNA247491.

AUTHOR CONTRIBUTIONS

DG and LA contributed conception and design of the study; DG selected the datasets and performed the computational analyses; DG and LA wrote the manuscript; LA edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets.2020.00365/full#supplementary-material

REFERENCES

1. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* (1999) 189:12–9. doi: 10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F

2. Jiang Z, Jhunjhunwala S, Liu J, Haverty PM, Kennemer MI, Guan Y, et al. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. *Genome Res.* (2012) 22:593–601. doi: 10.1101/gr.1339 26.111

3. Shannon-Lowe C, Rickinson AB, Bell AI. Epstein-Barr virus-associated lymphomas. *Philos Trans R Soc Lond B Biol Sci.* (2017) 372:20160271. doi: 10.1098/rstb.2016.0271

4. Hussein HAM, Okafor IB, Walker LR, Abdel-Raouf UM, Akula SM. Cellular and viral oncogenes: the key to unlocking unknowns of Kaposi's sarcoma-associated herpesvirus pathogenesis. *Arch Virol.* (2018) 163:2633–43. doi: 10.1007/s00705-018-3918-3

5. Derse D, Crise B, Li Y, Princler G, Lum N, Stewart C, et al. Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J Virol.* (2007) 81:6731–41. doi: 10.1128/JVI.02752-06

6. Liu W, MacDonald M, You J. Merkel cell polyomavirus infection and merkel cell carcinoma. *Curr Opin Virol.* (2016) 20:20–7. doi: 10.1016/j.coviro.2016.07.011

7. Chen Y, Williams V, Filippova M, Filippov V, Duerksen-Hughes P. Viral carcinogenesis: factors inducing DNA damage and virus integration. *Cancers.* (2014) 6:2155–86. doi: 10.3390/cancers 6042155

8. Sung W-K, Zheng H, Li S, Chen R, Liu X, Li Y, et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nature Genet.* (2012) 44:765–9. doi: 10.1038/ng.2295

9. Ayala-Diaz S, Jimenez-Lima R, Ramirez-Alcantara KM, Lizano M, Castro-Munoz LJ, Reyes-Hernandez DO, et al. Presence of papillomavirus DNA sequences in the canine transmissible venereal tumor (CTVT). *PeerJ.* (2019) 7:e7962. doi: 10.7717/peerj.7962

10. Luff J, Mader M, Rowland P, Britton M, Fass J, Yuan H. Viral genome integration of canine papillomavirus 16. *Papillomavirus Res.* (2019) 7:88–96. doi: 10.1016/j.pvr.2019.02.002

11. Munday JS, Tucker RS, Kiupel M, Harvey CJ. Multiple oral carcinomas associated with a novel papillomavirus in a dog. *J Vet Diagn Invest.* (2015) 27:221–5. doi: 10.1177/1040638714567191

12. Goldschmidt MH, Kennedy JS, Kennedy DR, Yuan H, Holt DE, Casal ML, et al. Severe papillomavirus infection progressing to metastatic squamous cell carcinoma in bone marrow-transplanted X-linked SCID dogs. *J Virol.* (2006) 80:6621–8. doi: 10.1128/JVI.02571-05

13. Priester WA, Mantel N. Occurrence of tumors in domestic animals. Data from 12 United States and Canadian colleges of veterinary medicine. *J Natl Cancer Inst.* (1971) 47:1333–44.

14. Beatty J. Viral causes of feline lymphoma: retroviruses and beyond. *Vet J.* (2014) 201:174–80. doi: 10.1016/j.tvjl.2014.05.026

15. Callanan JJ, Jones BA, Irvine J, Willett BJ, McCandlish IA, Jarrett O. Histologic classification and immunophenotype of lymphosarcomas in cats with naturally and experimentally acquired feline immunodeficiency virus infections. *Vet Pathol.* (1996) 33:264–72. doi: 10.1177/030098589603300302

16. Magden E, Miller C, MacMillan M, Bielefeldt-Ohmann H, Avery A, Quackenbush SL, et al. Acute virulent infection with feline immunodeficiency virus (FIV) results in lymphomagenesis via an indirect mechanism. *Virology.* (2013) 436:284–94. doi: 10.1016/j.virol.2012.12.003

17. Beatty JA, Callanan JJ, Terry A, Jarrett O, Neil JC. Molecular and immunophenotypical characterization of a feline immunodeficiency virus (FIV)-associated lymphoma: a direct role for FIV in B-lymphocyte transformation? *J Virol.* (1998) 72:767–71. doi: 10.1128/JVI.72.1.767-771.1998

18. Ding D, Lou X, Hua D, Yu W, Li L, Wang J, et al. Recurrent targeted genes of hepatitis B virus in the liver cancer genomes identified by a next-generation sequencing-based approach. *PLoS Genet.* (2012) 8:e1003065. doi: 10.1371/journal.pgen.1003065

19. Wang Q, Jia P, Zhao Z. VirusFinder: software for efficient and accurate detection of viruses and their integration sites in host genomes through next generation sequencing data. *PLoS One.* (2013) 8:e64465. doi: 10.1371/journal.pone.0064465

20. Xia Y, Liu Y, Deng M, Xi R. Detecting virus integration sites based on multiple related sequencing data by VirTect. *BMC Med Genomics.* (2019) 12(Suppl 1):19. doi: 10.1186/s12920-018-0461-8

21. Chen Y, Yao H, Thompson EJ, Tannir NM, Weinstein JN, Su X. VirusSeq: software to identify viruses and their integration sites using next-generation sequencing of human cancer tissue. *Bioinformatics.* (2013) 29:266–7. doi: 10.1093/bioinformatics/bts665

22. Li JW, Wan R, Yu CS, Co NN, Wong N, Chan TF. ViralFusionSeq: accurately discover viral integration events and reconstruct fusion transcripts at single-base resolution. *Bioinformatics.* (2013) 29:649–51. doi: 10.1093/bioinformatics/btt011

23. Ho DW, Sze KM, Ng IO. Virus-Clip: a fast and memory-efficient viral integration site detection tool at single-base resolution with annotation capability. *Oncotarget.* (2015) 6:20959–63. doi: 10.18632/oncotarget.4187

24. Wang Q, Jia P, Zhao Z. VERSE: a novel approach to detect virus integration in host genomes through reference genome customization. *Genome Med.* (2015) 7:2. doi: 10.1186/s13073-015-0126-6

25. Elvers I, Turner-Maier J, Swofford R, Koltookian M, Johnson J, Stewart C, et al. Exome sequencing of lymphomas from three dog breeds reveals somatic mutation patterns reflecting genetic background. *Genome Res.* (2015) 25:1634–45. doi: 10.1101/gr.194449.115

26. Ramsey SA, Xu T, Goodall C, Rhodes AC, Kashyap A, He J, et al. Cross-species analysis of the canine and human bladder cancer transcriptome and exome. *Genes Chromosomes Cancer.* (2017) 56:328–43. doi: 10.1002/gcc.22441

27. Wang G, Wu M, Maloneyhuss MA, Wojcik J, Durham AC, Mason NJ, et al. Actionable mutations in canine hemangiosarcoma. *PLoS One.* (2017) 12:e0188667. doi: 10.1371/journal.pone.0188667

28. Hendricks WPD, Zismann V, Sivaprakasam K, Legendre C, Poorman K, Tembe W, et al. Somatic inactivating PTPRJ mutations and dysregulated pathways identified in canine malignant melanoma by integrated comparative genomic analysis. *PLoS Genet.* (2018) 14:e1007589. doi: 10.1371/journal.pgen.1007589

29. Saffari PS, Vapniarsky N, Pollack AS, Gong X, Vennam S, Pollack AJ, et al. Most canine ameloblastomas harbor HRAS mutations, providing a novel large-animal model of RAS-driven cancer. *Oncogenesis.* (2019) 8:11. doi: 10.1038/s41389-019-0119-1

30. Gardner HL, Sivaprakasam K, Briones N, Zismann V, Perdigones N, Drenner K, et al. Canine osteosarcoma genome sequencing identifies recurrent mutations in DMD and the histone methyltransferase gene SETD2. *Commun Biol.* (2019) 2:266. doi: 10.1038/s42003-019-0487-2

31. Amin SB, Anderson KJ, Boudreau CE, Martinez-Ledesma E, Kocakavuk E, Johnson KC, et al. Comparative molecular life history of spontaneous canine and human gliomas. *Cancer Cell.* (2020) 37:243–57.e7. doi: 10.1016/j.ccell.2020.01.004

32. Lee KH, Hwang HJ, Noh HJ, Shin TJ, Cho JY. Somatic mutation of PIK3CA (H1047R) is a common driver mutation hotspot in canine mammary tumors as well as human breast cancers. *Cancers.* (2019) 11:2006. doi: 10.3390/cancers11122006

33. Vozdova M, Kubickova S, Pal K, Fröhlich J, Fictum P, Rubes J. Recurrent gene mutations detected in canine mast cell tumours by next generation sequencing. *Vet Comp Oncol.* (2020). doi: 10.1111/vco.12572. [Epub ahead of print].

34. Chen X, Kost J, Li D. Comprehensive comparative analysis of methods and software for identifying viral integrations. *Brief Bioinform.* (2018) 20:2088–2097. doi: 10.1093/bib/bby070

35. Bhaduri A, Qu K, Lee CS, Ungewickell A, Khavari PA. Rapid identification of non-human sequences in high-throughput sequencing datasets. *Bioinformatics.* (2012) 28:1174–5. doi: 10.1093/bioinformatics/bts100

36. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank. *Nucleic Acids Res.* (2008) 36(Database issue):D25–30. doi: 10.1093/nar/gkm929

37. Cantalupo PG, Calgua B, Zhao G, Hundesa A, Wier AD, Katz JP, et al. Raw sewage harbors diverse viral populations. *mBio.* (2011) 2:e00180–11. doi: 10.1128/mBio.00180-11

38. Tang KW, Alaei-Mahabadi B, Samuelsson T, Lindh M, Larsson E. The landscape of viral expression and host gene fusion and adaptation in human cancer. *Nature Commun.* (2013) 4:2513. doi: 10.1038/ncomms3513

39. Cantalupo PG, Pipas JM. Detecting viral sequences in NGS data. *Curr Opin Virol.* (2019) 39:41–8. doi: 10.1016/j.coviro.2019.07.010

40. Gompels UA, Nicholas J, Lawrence G, Jones M, Thomson BJ, Martin ME, et al. The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. *Virology.* (1995) 209:29–51. doi: 10.1006/viro.1995.1228

41. Fick LJ, Fick GH, Li Z, Cao E, Bao B, Heffelfinger D, et al. Telomere length correlates with life span of dog breeds. *Cell Rep.* (2012) 2:1530–6. doi: 10.1016/j.celrep.2012.11.021

42. Forster M, Szymczak S, Ellinghaus D, Hemmrich G, Ruhlemann M, Kraemer L, et al. Vy-PER: eliminating false positive detection of virus integration events in next generation sequencing data. *Sci Rep.* (2015) 5:11534. doi: 10.1038/srep11534

43. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* (2001) 409:860–921. doi: 10.1038/35057062

44. Boeke JD, Stoye JP. Retrotransposons, Endogenous Retroviruses, and the Evolution of Retroelements. In: Coffin JM, Hughes SH, Varmus HE, editors. *Retroviruses.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1997). p. 343–435.

45. Roy-Burman P. Endogenous env elements: partners in generation of pathogenic feline leukemia viruses. *Virus Genes.* (1995) 11:147–61. doi: 10.1007/BF01728655

46. Halo JV, Pendleton AL, Jarosz AS, Gifford RJ, Day ML, Kidd JM. Origin and recent expansion of an endogenous gammaretroviral lineage in domestic and wild canids. *Retrovirology.* (2019) 16:6. doi: 10.1186/s12977-019-0468-z

47. Martínez Barrio Á, Ekerljung M, Jern P, Benachenhou F, Sperber GO, Bongcam-Rudloff E, et al. The first sequenced carnivore genome shows complex host-endogenous retrovirus relationships. *PLoS One.* (2011) 6:e19832. doi: 10.1371/journal.pone.0019832

48. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature.* (2005) 438:803–19. doi: 10.1038/nature04338

49. Modiano JF, Getzy DM, Akol KG, Van Winkle TJ, Cockerell GL. Retrovirus-like activity in an immunosuppressed dog: pathological and immunological findings. *J Comp Pathol.* (1995) 112:165–83. doi: 10.1016/S0021-9975(05)80059-3

50. Ghernati I, Auger C, Chabanne L, Corbin A, Bonnefont C, Magnol JP, et al. Characterization of a canine long-term T cell line (DLC 01) established from a dog with Sezary syndrome and producing retroviral particles. *Leukemia.* (1999) 13:1281–90. doi: 10.1038/sj.leu.2401480

51. Fadel HJ, Poeschla EM. Retroviral restriction and dependency factors in primates and carnivores. *Vet Immunol Immunopathol.* (2011) 143:179–89. doi: 10.1016/j.vetimm.2011.06.002

52. Tarlinton RE, Barfoot HK, Allen CE, Brown K, Gifford RJ, Emes RD. Characterisation of a group of endogenous gammaretroviruses in the canine genome. *Vet J.* (2013) 196:28–33. doi: 10.1016/j.tvjl.2012.08.011

53. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* (2004) 2:E234. doi: 10.1371/journal.pbio.0020234

54. Bodelon C, Untereiner ME, Machiela MJ, Vinokurova S, Wentzensen N. Genomic characterization of viral integration sites in HPV-related cancers. *Int J Cancer.* (2016) 139:2001–11. doi: 10.1002/ijc.30243

55. Budzinska MA, Shackel NA, Urban S, Tu T. Cellular genomic sites of hepatitis B virus DNA integration. *Genes (Basel).* (2018) 9:365. doi: 10.3390/genes9070365

56. Aresu L, Ferraresso S, Marconato L, Cascione L, Napoli S, Gaudio E, et al. New molecular and therapeutic insights into canine diffuse large B cell lymphoma elucidates the role of the dog as a model for human disease. *Haematologica.* (2018) 104:e256–e259. doi: 10.3324/haematol.2018.207027

57. Chu S, Wylie TN, Wylie KM, Johnson GC, Skidmore ZL, Fleer M, et al. A virome sequencing approach to feline oral squamous cell carcinoma to evaluate viral causative factors. *Vet Microbiol.* (2020) 240:108491. doi: 10.1016/j.vetmic.2019.108491